THE BRITISH JOURNAL OF EDUCATIONAL PSYCHOLOGY

24

FEBRUARY 1985

VOL. 55, PART 1

THE
BRITISH JOURNAL
OF
EDUCATIONAL PSYCHOLOGY

Edited by
Professor GEOFFREY BROWN

Assistant Editors:
Dr. R. L. BURDEN
Dr. C. W. DESFORGES
Dr. D. FONTANA

Sub-Editor:
Dr. S. CUNNINGHAM

Book Review Editor:
Dr. C. W. DESFORGES

Advised by the Editorial Committee:
Dr. H. BELOFF
Professor M. B. SUTHERLAND
Professor J. WRIGLEY

Vol. 55, Part 1

The British Journal of Educational Psychology is issued by the British Psychological Society once a term, in February, June and November. The subscription is £18·00 a year post free, payable in advance, or £6·50 each number.

Subscriptions and Orders should be sent to Scottish Academic Press, 33 Montgomery Street, Edinburgh, EH7 5JX. Members of the British Psychological Society receive the *Journal* on special terms.

Papers for publication in duplicate should be sent to Professor Geoffrey Brown, School of Education, University of East Anglia, Norwich NR4 7TJ. Guidance to contributors on the preparation of papers for the *Journal* is given at the end of the November number each year.

All *Books for Review* for *The British Journal of Educational Psychology* should be sent to: Dr. C. W. Desforges, School of Education, University of East Anglia, Norwich NR4 7TJ.

Advertisements for this *Journal* should be sent to Scottish Academic Press.

Published by
SCOTTISH ACADEMIC PRESS, 33 MONTGOMERY STREET, EDINBURGH, EH7 5JX,
for the British Journal of Educational Psychology, Ltd.

All rights reserved

© 1985. The British Journal of Educational Psychology

PRINTED IN GREAT BRITAIN BY CLARK CONSTABLE, EDINBURGH, LONDON, MELBOURNE

P3889

Br. J. educ. Psychol., 55, 1-10, 1985

THE INFLUENCE OF INFORMATION AND DAILY CONTACT ON CHILDREN'S ATTITUDES TOWARDS APHASIC CHILDREN

24

By DAN ZAKAY

(*Department of Psychology, Tel-Aviv University, Israel*)

SUMMARY. The influences upon attitudes towards aphasic children were studied when ordinary children (ages 9-14) were given explanations of, and contact with, children with motoric aphasia. Subjects were 191 elementary school pupils. In one group no explanation about aphasia was given, and no contact with aphasic children was maintained. In a second group, an explanation was given but no contact was maintained. In the third group, a special class of aphasic children was incorporated within the ordinary school. As a result, daily contact between the ordinary and the aphasic children was maintained. A questionnaire for measuring attitudes towards aphasic children was built and its test-retest reliability was found to be $0{\cdot}91$ $(\mathrm{df}=37, P<0{\cdot}01)$. The attitude scores were analysed by analysis of variance. A significant main effect of treatment method was found $(F\,(2,173)=5{\cdot}38, P<0{\cdot}05)$. As hypothesised, the attitudes of children exposed to daily contact were less negative than those of children exposed to explanation only. The attitudes of children who were exposed neither to explanation nor to contact were the most negative. A significant main effect of age was found $(F\,(2,173)=10{\cdot}84, P<0{\cdot}01)$. It was revealed, as hypothesised, that attitudes towards aphasic children were more negative in younger children than older ones, and only the attitudes of the 13- to 14-year-old children were in the positive region. A significant interaction between age and treatment method was also found $(F\,(4,173)=3{\cdot}81, P<0{\cdot}05)$. The contact was found to be most effective with the youngest children. The implications of the findings for the general question of attitude formation in children and for the specific question of aphasic children's rehabilitation were discussed.

INTRODUCTION

THE present research deals with the attitudes of normal children towards exceptional children, specifically children diagnosed as suffering from motoric aphasia. Such children are characterised as having a level of intelligence that is within the normal range for their age (Barry, 1961; Meginnis, 1963). The difficulties caused by the motoric aphasia are to be found in the child's inability to repeat words and sounds, as well as various memory impairments, an inability to concentrate, hyperactivity and impulsivity, emotional instability, slight motoric difficulties mainly in activities demanding co-ordination, and, on occasion, hearing impairments. The facial appearance of aphasic children is normal and many of them might be described as being handsome. However, their exceptional external behaviour and their difficulties in talking and communicating expose them to social rejection, prejudices, and stereotyping by normal children.

The formation of negative attitudes by normal children towards exceptional children is not surprising in light of psychological knowledge:

(1) Cruickshank (1966) notes that physical deviations are perceived as undesirable, even in cases where they do not impose any restrictions on abilities needed in order to function normally in the environment. Wright (1960) claims that any physical impairment is perceived as loss of a valued asset.

(2) Exceptional children, as any exceptional group, are probably categorised as belonging to some minority group (Wright, 1960) and, hence, the formation of negative attitudes towards exceptional children might be similar to the formation of negative attitudes towards minority groups. This explanation is supported by the

P 3889

finding (Ehrlich, 1973) that negative attitudes towards the blind were found to be significantly correlated with negative attitudes towards minority groups.

(3) Wright suggests that a further reason for such a negative viewpoint is the common belief that exceptional children tend to become a heavy burden for others.

(4) Another factor suggested by Wright is that illness is perceived by children as a form of punishment, and is associated with not keeping hygienic rules. Indeed, children tend to associate illness with sin.

(5) Any encounter with exceptional individuals is basically an unfamiliar and unstructured situation that tends to be a threatening one — and thus one which most people prefer to avoid (Heider, 1958). It is of interest to note that the above-mentioned factors coincide with the factors noted by Jahoda (1971) as being possible causes for the formation of superstitions.

(6) Cruickshank (1963) reported that the children most popular among their peers were those who were successful in quickly adjusting and adapting themselves to new situations. As the probability for an exceptional child to adjust quickly to new situations is, by definition, lower than that of an ordinary child, the probability of him/her being popular and socially accepted is also lower.

The negative attitudes towards exceptional children and their social rejection have negative implications regarding their chances of rehabilitation. Anastasi (1979) noted that "handicapped persons of any age tend to have special emotional problems, such as feelings of insecurity or hostility and various types of defensive reactions. These reactions often stem from a feeling of being different from one's associates" (p. 401). Anastasi further claimed that in the case of children, delayed psychological problems might develop as a result of the curtailment of normal experiences. Cruickshank (1966) noted that because of social rejection, the life space of exceptional children is restricted and less complex than that of normal children, and their self-image is harmed.

Exceptional children find themselves trapped within a vicious circle. Because of social rejection they develop more antisocial habits such as withdrawal and shyness which, in turn, strengthen that social rejection. Hence, it seems to be of great importance to create a positive, or at least a neutral, atmosphere around exceptional children. In order to achieve such a goal, methods for changing negative attitudes and/or preventing their formation must be developed. As such, we are dealing with a non-congruent attitude change, which is far more difficult to achieve than a congruent one (Zajonc, 1960; McGuire, 1962).

It is commonly held (e.g., Cialdini *et al.*, 1981) that the three most effective methods for causing a non-congruent attitude change are: (1) supplying information; (2) a change in group affiliation; and (3) interpersonal contact between those who hold the attitude to be changed, and those who serve as referent for that attitude. The positive interpersonal contact with a source of positive information is considered to be the most effective means of attitude change with children (Rokeach, 1961; Thompson, 1962). This is also in accordance with Allport's (1954) contact theory, which was supported in the frame of racial attitudes (e.g., Cook, 1969; Schofield and Sugar, 1977). Indeed, Guralnick (1976) reported an improvement in social behaviour of preschool handicapped children as a result of integration with non-handicapped peers. Similar results were reported by Young and Keri (1975) regarding the behaviour of severely retarded schoolchildren.

However, these studies dealt with the effect of the contact on the disabled children, and did not test whether any corresponding attitude change occurred among the normal children. This question was raised by Handlers and Austin (1980) with regard to the attitudes of high school students towards their handicapped peers. The findings were that, as a result of integration between handicapped and non-handicapped students, 82 per cent of the non-handicapped subjects felt a positive change in their attitude, and 62 per cent felt that direct interpersonal contact was the most effective means for ensuring such a positive shift. However, these results cannot be generalised to elementary school children. Furthermore, the critical period for the development of a stable positive attitude towards disabled children as well as for the creation of a positive effect on the disabled child is at the pre-school and school-age. Thus, it is important to test the influence of information and direct interpersonal contact on attitudes towards disabled children in these years.

This study deals with the aforementioned question of attitudes towards aphasic children within the framework of schoolchildren. The following hypotheses were tested:

(1) The attitude of normal schoolchildren towards aphasic children will be negative, with an intensity that decreases with age.

(2) The attitude of normal schoolchildren towards aphasic children will be significantly more positive among those who have had direct contact and information *vis-à-vis* aphasic children, as compared with that of normal children who have received only information about aphasics, and to those who have neither been exposed to information nor had direct contact with aphasic children.

METHOD

Subjects

Ninety-six boys and 95 girls from three elementary schools (in Tel Aviv), matched for socio-economic level, participated in the study. Sixty-one were 4th-grade pupils (ages 9-10), 62 were 6th-graders (ages 11-12), and 68 were 8th-graders (ages 13-14). In addition, 213 children who did not participate in the main study were chosen in equal number from each grade in each school for the instrument-building phase of the study. The study was conducted with the complete agreement and participation of teachers and school administrators.

The experimental school

One of the three schools was chosen by the municipality as the subject of an experimental programme according to which a special class, consisting of five 8-year-old aphasic children (four boys and one girl), was integrated into the regular school life. While this group comprised only 1 per cent of the school population, there was nevertheless no doubt that all of that population did come into some contact with the aphasic children. This special class of five was incorporated within the school in a series of steps. *Stage 1*: Two pleasant rooms in a separate wing were allocated, one serving as classroom, and the other as nurse's room. The regular 1st grade classroom was located in the vicinity. *Stage 2*: While the class was being set up in the school, a series of preparatory talks was presented to the regular classes. The purpose was twofold: to explain the issue, and to generate a conducive atmosphere. *Stage 3*: A month after the aphasic children had entered the school, they were taken out into the schoolyard during recess, under the supervision of most of the teachers. Within the second month of their being in the school, social integration gradually developed. The aphasic children began to participate in group activities, and a number of children began to contact the aphasic children.

Control groups

As the study commenced at a stage where the aphasic children had already been integrated within the experimental school, there was no possibility of conducting a pre-post designed study. For this reason, two additional schools were selected to serve as matched controls for the present experimental school. For these controls two experimental conditions were manipulated: (1) No reference and acquaintance with the problem of aphasic children (NIC group); (2) Acquaintance with the problem and receiving information about aphasic children (identical to Stage 2 in the experimental school), but no direct contact with them (INC group).

Instruments

A. Attitude questionnaire — A sample of 213 children were requested to write down five positive and five negative traits for elementary school children. All in all, a sum of 357 traits were provided by the children, 152 of which were positive, and 205 negative. The questionnaire (available from the author) was composed of the 10 most frequent positive and the 10 most frequent negative traits in random order. Attitudes were measured in the following way: two descriptions of children were constructed, one conceivably pertaining to the description of an aphasic child (Child A), and the second to an ordinary child (Child B). The descriptions were validated by the teachers of the aphasic class as representative. For each trait, subjects were requested to indicate, given the two descriptions, whether it was appropriate to assign the trait to each of the described children. Each of the described children was referred to separately, in terms of three categories: "Yes", "Don't Know", and "No".

Scoring method — The scoring method used is presented in Table 1. This method expresses the comparative relation between the measure of match of each child to each trait. If both descriptions are matched equally, a neutral score of zero is given. The overall score is defined by the sum of scores of all 20 traits, and its range is between +20 and −20. The reliability of the questionnaire was tested by a test-retest and was found to be 0·91 (df=37; P<0·01). The questionnaire was validated in the following way: a number of pupils from the experimental school, who showed distinctive feelings (either positive or negative) towards the aphasic children, were identified and the questionnaire was administered to them. In all cases it was found that positive and negative feelings were accompanied by positive or negative attitude scores, respectively.

TABLE 1

SCORING METHOD IN THE ATTITUDE QUESTIONNAIRE

Child B Response	Child A Response		
	Yes	No	Don't know
Yes	0	−1*	−0·5*
No	+1*	0	+0·5*
Don't know	+0·5*	−0·5*	0

* This score is for positive traits and the sign is reversed for negative traits.

B. Preference measurement — In order to test for the willingness of chidren to come into direct contact with aphasic children, two situations were selected: friendship and membership in the class committee. Each subject was asked: (1) Who would you prefer to have as your friend — Child A or Child B? (2) Who would you like to have as member of the class committee — Child A or Child B? Regarding friendship, subjects were requested to detail their answer.

Procedure

In school E (experimental) and in school NIC (no information and no contact) the questionnaires were administered with no prior preparation. In school INC (information, no contact) an identical process to that which took place in school E (Stage 2) was carried out a week prior to the presentation of the questionnaires. The questionnaire was administered in the same manner to all schools. Attempts were made to ensure that all children clearly understood their task. The children were told that there were no correct or incorrect answers and that the required answer was that which they felt was appropriate. It was emphasised that the questionnaires had no relevance to studies, grades, or school certificates. No time limit was imposed. All questionnaires were administered by class tutors within the same week, at a period three months after the start of the experimental programme in the experimental school.

RESULTS

Attitudes

The attitude score of each child *vis-à-vis* the aphasic child was calculated. The mean attitude scores according to type of treatment and grade are presented in Table 2. These attitude scores were analysed by means of a three-way complete factorial ANOVA with treatment type, grade, and sex being the three independent variables. Significant main effects of type of treatment (F (2,173)=5·38; P<0·05) and grade (F (2,173)=10·845; P<0·01) were obtained, while the main effect of sex was not found to be significant. Moreover, a significant interaction between treatment type and grade (F (4,173)=3·81; P<0·05) was found. This interaction is presented in Figure 1.

TABLE 2

ATTITUDE MEANS TOWARDS THE APHASIC CHILDREN
(N is given in parentheses)

	Grade			
	4th	6th	8th	Total
Group				
E	−2·65 (20)	−2·95 (22)	10·83 (24)	−1·39 (66)
INC	−5·80 (21)	−4·22 (20)	−0·71 (21)	−3·35 (62)
NIC	−16·87 (20)	−2·00 (20)	−2·63 (23)	−6·95 (63)
Total	−8·40 (61)	−3·05 (62)	1·29 (68)	−3·86 (191)

Specific comparisons were conducted on the attitude scores using Schéffé tests. The following comparisons yielded significant results:

(*a*) Group E *versus* Group NIC (F (2,188)=5·33; P<0·01);
(*b*) Grade 4 *versus* Grade 6 (F (2,188)=4·70; P<0·01);
(*c*) Grade 4 *versus* Grade 8 (F (2,188)=16·18; P<0·01);
(*d*) Grade 6 *versus* Grade 8 (F (2,188)=3·28; P<0·05);
(*e*) Group E *versus* Group NIC within the 4th grade (F (8,183)=5·10; P<0·01).

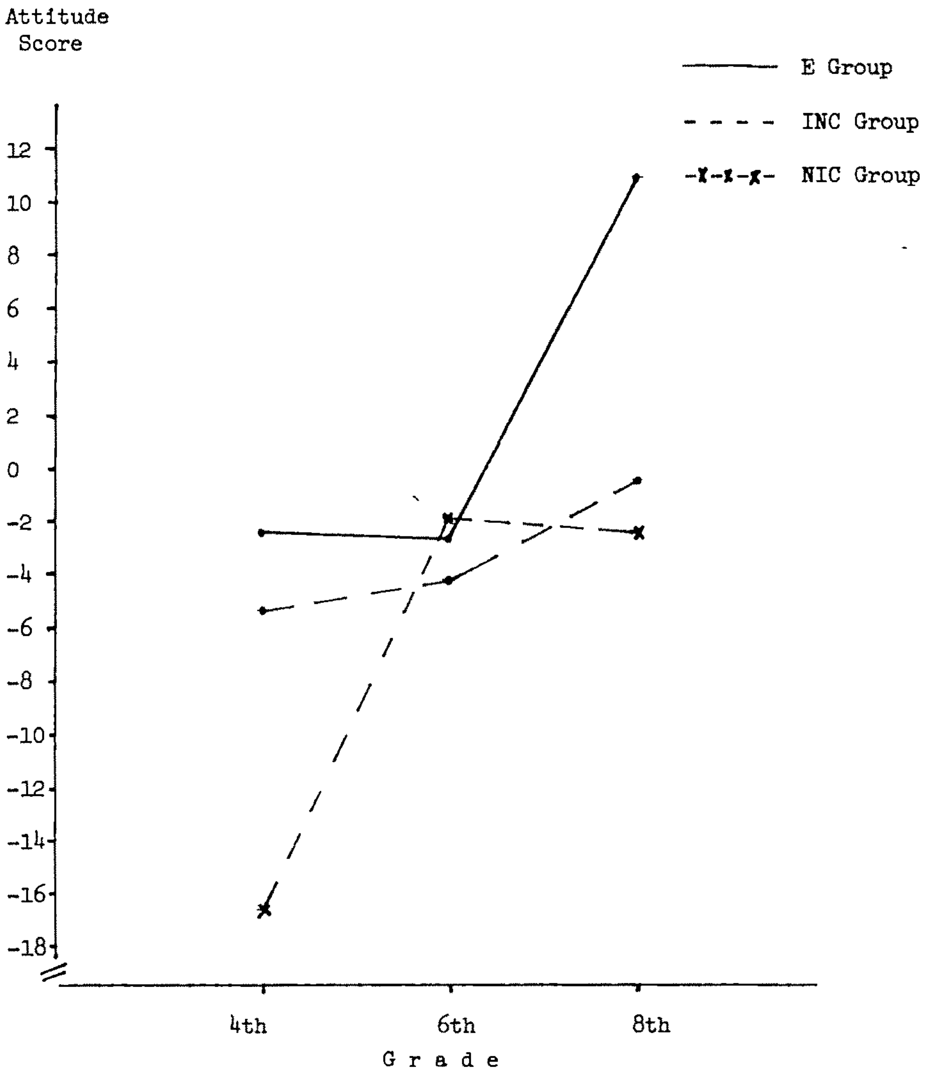
All of the results are in the hypothesised direction.

Preference test

The proportion of children who preferred the aphasic child over the ordinary child in each situation was calculated. In order more fully to understand these preferences regarding friendship, a content analysis was performed on the explanations given by the children for that preference. The explanations were categorised as belonging to one of two categories: preference based on compassion, or preference for standard reasons of friendship. This content analysis was conducted by two independent judges and the

FIGURE 1

THE INTERACTION BETWEEN TYPE OF TREATMENT AND AGE



agreement between them was complete. The proportions of the various preferences according to grades and types of treatment are presented in Table 3 and in Figure 2.

The proportions were analysed by a series of chi square tests. The following comparisons were found to be significant:

(*a*) Choosing the aphasic child as a friend was significantly higher in Group E as compared with Group NIC (χ^2 (df = 1) = 4·812; $P < 0{\cdot}05$). Furthermore, this finding holds for the 4th grade by itself (χ^2 (df = 1) = 8·18; $P < 0{\cdot}01$), while in the 8th grade this is true only when Group E was compared with Group INC (χ^2 (df = 1) = 5·23; $P < 0{\cdot}01$).

(*b*) No significant differences were found regarding preferences for class committee member. Almost all respondents preferred the ordinary child in this respect.

(*c*) The proportion of children who were willing to choose the aphasic child as a friend for regular reasons was significantly higher in Group E as compared with Group NIC (χ^2 (df=1)=5·54; $P<0{\cdot}05$). This is true also when Grade 8 was compared with Grade 6 (χ^2 (df=1)=3·88; $P<0{\cdot}05$) across type of treatments. However, when each treatment group was analysed according to age, the only significant difference was found within the 4th grade, between Groups E and NIC (Fisher test; $P<0{\cdot}05$).

TABLE 3

APHASIC CHILD'S PREFERENCES PROPORTIONS

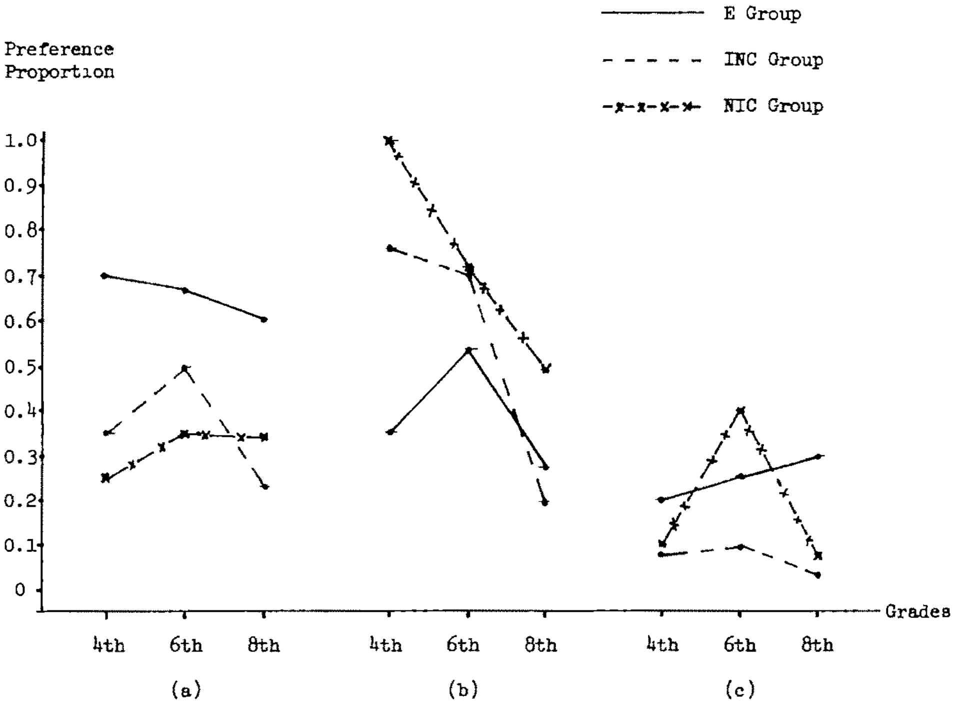
	As friend			As friend out of compassion*			As committee member		
Grade.	4th	6th	8th	4th	6th	8th	4th	6th	8th
Group									
E	0·70	0·68	0·60	0·35	0·53	0·28	0·20	0·27	0·30
INC	0·36	0·50	0·22	0·75	0·70	0·20	0·09	0·10	0·04
NIC	0·25	0·35	0·34	1·00	0·71	0·50	0·10	0·40	0·08

* Out of those preferring the aphasic child as a friend.

FIGURE 2

APHASIC CHILD'S PREFERENCE'S PROPORTIONS·

(*a*) as a friend, (*b*) as a friend out of compassion, (*c*) as class committee member



Nevertheless, the significance of the finding of a reduction in negative attitude towards aphasic children should not be diminished. A follow-up study revealed that a year later there was a marked positive improvement in the condition and adaptation of the aphasic children. Thus, there is a definite need to continue in these efforts. The contribution is both to the ordinary children as well as to the aphasics.

The present study suggests that it is not sufficient to rely on information only, and that rather interpersonal contact has to be encouraged. Even this, however, is not sufficient, and one has to look for new and more effective means. Thus, for example, Lancioni (1982) found that normal children could successfully pass social conduct on to their retarded peers. It would appear that there is room for an approach that requires intimate contact between the normal and exceptional child, in which a certain responsibility is placed on the normal child. It might be that such an approach would be most effective for inducing attitude change on behalf of normal children. It is also of interest to consider new means to allow for attitude change in the younger children. Conceivably, the necessary means here would involve the parents, and attempts might even be made to induce attitude change in them so that their children would incorporate such a shift, being directly influenced by their major information source — their own parents.

Perhaps one may explain the importance of contact on one's attitude in terms of script theory (Abelson, 1981), which views social behaviour as being determined by the application of a certain script. A positive contact enables the generation of new scripts regarding prior negative attitudes, which in turn increases the probability of eliciting positive behaviour in subsequent contacts. It is possible that theoretical explanations are not, by themselves, sufficient to alter old scripts or create new ones. The importance of the cognitive factor in the process of attitude change formation is emphasised in the review of the literature by Cialdino *et al.* (1981), and it would seem that it is important to incorporate this channel in attempts to alter attitudes among children. Needless to say, all of the suggestions made should be tested in further studies, to enable empirical validation.

REFERENCES

Abelson, R. P. (1981). Script theory. *Am Psychol.*, 36, 715-729.

Allport, G. W. (1954). *The Nature of Prejudice.* Reading, Mass. Addison-Wesley.

Anastasi, A (1979). *Fields of Applied Psychology.* New York: McGraw-Hill.

Barry, H. (1961). *The Young Aphasic Child.* New York: Bell Association.

Cialdini, R. B., Petty, R. E. and Cacioppa, J. T. (1981). Attitude and attitude change. *Ann. Rev. Psychol.*, 32, 357-404.

Cook, S. W (1969). Motives in a conceptual analysis of attitude-related behavior. In Arnold, W. J. and Levine, D. (Eds.), *Nebraska Symposium on Motivation* (Vol. 17), Lincoln: University of Nebraska Press.

Cook, S. W. (1971). The effect of unintended interracial contact upon racial interaction and attitude change. *Project No. 5-1320. Final Report.* U.S. Office of Education.

Cruickshank, W. M. (1963). *Psychology of Exceptional Children and Youth.* Englewood Cliffs, N.J : Prentice-Hall.

Cruickshank, W M (1966). *The Teacher of Brain-injured Children.* Syracuse: Syracuse University Press.

Ehrlich, H. K. (1973). *The Social Psychology of Prejudice.* New York Wiley.

Guralnick, M. J. (1976). The value of integrating handicapped and nonhandicapped pre-school children. *Am. J. Orthopsychiatry*, 42, 236-245.

Handlers, A., and Anstin, K. (1980). Improving attitudes of high school students toward their handicapped peers. *Except. Children*, 47, 228-229.

Heider, F (1958) *The Psychology of Interpersonal Relations.* New York· Wiley.

Jahoda, G. (1971). *The Psychology of Superstition.* New York: Penguin Books.

Lambert, W. E., and Klineberg, O. (1967). *Children's Views of Foreign Peoples.* New York: Appleton-Century-Crofts.

Lancioni, G. E. (1982). Normal children as tutors to teach social responses to withdrawn mentally retarded schoolmates· training, maintenance and generalization. *J. appl. Behav. Analysis*, 15, 17-40.

McGuire, W. J. (1962). Persistence of the resistance to persuasion induced by various types of prior belief defenses. *J. abnorm. soc. Psychol.*, 64, 241-248.
Meginnis, M. A. (1963). *Aphasic Children*. New York: Bell Association.
Schofield, J. W., and Sugar, H. A. (1977). Peer interaction pattern in an integrated middle school. *Sociometry*, 40, 130-138.
Rokeach, M. (1961). *The Open and Closed Mind*. New York. Basic Books.
Thompson, G. G. (1962). *Child Psychology*. New York: Houghton-Mifflin.
Wright, B. A. (1960) *Physical Disability — A Psychological Approach*. New York: Harper Row.
Young, C. C., and Kerr, M. M (1975). The effects of a retarded child's social initiations on the behavior of severely retarded school aged peers. *Education and Training of the Mentally Retarded*, 14, 185-190.
Zajonc, R. B. (1960). The concepts of balance, congruity and dissonance. *Public Opinion Q.*, 24, 280-296.

(*Manuscript received* 4*th July*, 1984)

Br. J. educ. Psychol., 55, 11-23, 1985

THE CHILD'S UNDERSTANDING OF THE CAUSES OF ACADEMIC SUCCESS AND FAILURE: A CASE STUDY OF BRITISH SCHOOLCHILDREN.

By A. W. LITTLE

(Institute of Development Studies at the University of Sussex)

SUMMARY. Current models of the attribution process in children derive from adult models. This study examines the adequacy of current models for British children in the 5-14 age range (N = 149). Three questions are asked: (1) What types of explanation are used freely by children to explain academic success and failure? (2) How frequently are they used? (3) Do the types of explanation used by children to explain academic success and failure vary developmentally? Eighteen categories are used to classify the free responses of children to six achievement stories. The categories used by current models are useful but limited. Additional categories include different types of ability attribution and teacher-sanctioned behaviour. Reservations are expressed about the meaning and relative importance of luck and the organising dimensions (locus, stability, control) of current taxonomical models. Using a multivariate profile analysis it was shown that the pattern of use of attribution categories varied by age. While recognising the importance of situational explanations for this trend, it was concluded that the trend was not reducible to them and that the attribution process involves a complex interaction of subjective and objective reality.

INTRODUCTION

THE number of studies of children's attributions for success and failure has increased substantially over the past decade. In 1976 Weiner and Kun were able to provide an exhaustive survey of work on developmental aspects of attribution during the sixties and early seventies (Weiner and Kun, 1976). Since then a number of review articles have appeared each focusing on and updating different sets of developmental attribution literature (e.g., Weiner, 1979; Weiner *et al.*, 1980; Kelley and Michela, 1980; Frieze, 1980; Fincham, 1982). The vast majority of studies reviewed use Weiner's early (Weiner *et al.*, 1971) or later (Weiner, 1979) taxonomical model of attributions. The 1971 model is a 2 × 2 matrix in which ability, effort, task difficulty and luck are organised along the dimensions of locus and stability. This model is an anchor point for research questions about the antecedents and consequences of ability, effort, task difficulty and luck attributions. The more recent model incorporates these and a few additional attributions but uses three rather than two organising dimensions — locus, stability and control.

However, it is important to note that the models were developed to describe adult attributions. Furthermore they assume that the range and type of children's attributions are the same as those of adults; that the dimensions used to organise attributions are the same for children of all ages and adults; that all attributions conform to a causal structure of efficient causality; and that the meanings attached to attributions are the same for children and adults. The depth of these assumptions is perhaps better appreciated when one realises that even adults' attributions, on which attribution theories are constructed, have been derived from the intuitions of the attribution theorists (e.g., Weiner, 1979, p. 4). Only quite recently did Kelley and Michela (1980) suggest that the central irony of attribution research was the lack of research into the causal distinctions made by common people. This paper examines the adequacy of current models of attribution for British children in the 5-14 age range. It seeks to answer three questions: (1) What types of explanations are used freely by children to explain academic success and failure; (2) How frequently are they used? (3) Do the types of explanation used by children to explain academic success and failure vary developmentally?

METHOD

The sample consisted of 149 schoolchildren in the age range 5 yrs. 2 m.-14 yrs. 8 m. divided into four age groupings from one primary school and one secondary school in the county of Sussex. The secondary schoolchildren had attended the sampled primary school. Mean achievement and social class levels were average for the county and the country. Details of sample size, age groupings, sex distribution and mean deviation IQ score (Otis-Lennon) are shown below in Table 1.

TABLE 1

MEAN AGE, AGE RANGE, SAMPLE SIZE, SEX DISTRIBUTION AND MEAN DEVIATION IQ SCORE

	Age group			
	1	2	3	4
Mean age	5 yr. 7 m.	8 yr. 1 m.	11 yr. 3 m.	14 yr. 2 m.
Age range	(5·2-6·0)	(7·7-8·7)	(10·9-11·9)	(13·9-14·8)
N	34	37	36	42
% Female	50	51	53	57
Mean deviation IQ score	102	108	100	103

Interviews were conducted individually and tape-recorded by the author. The interview consisted of six simple stories* representing a range of achievement events typically found in British classrooms. They were:

Story 1. Do you think that anyone in your class is working well? *Can you think of the reason why X is working well?* Can you think of any other reasons?
Story 2. Do you think that anyone in your class is working not so well? *Can you think of the reason why Y is not working very well?* Can you think of any other reasons?
Story 3. Do you think that anyone in your class is working better than anyone else? Who are they working better than? *What is the reason why X is working better than Y?* Can you think of any other reasons?
Story 4. Let's think about two children who are working on their reading books. One child has finished reading book 6 and the other child has finished reading book 5. *Can you think of the reason why one child has finished six reading books and one has finished five?* Can you think of any other reasons?
Story 5. Let's think about another two children who are working on their number work. One child has finished his Fletcher book and the other child has finished half of the same book. *Can you think of the reason why one child has finished all of the book and one has finished half of the book?* Can you think of any other reasons?
Story 6. Let's think about another two children who are in an Art lesson. One child has painted a very nice painting. The other child has painted not such a nice painting. *Can you think of the reason why one child has done a very nice painting and the other child has done not such a nice painting?* Can you think of any other reasons?

The key causal question is in italic. Children were encouraged to respond in whatever way they chose and to produce as many causal statements as possible.† Taped interviews were transcribed by the author into extensive sets of verbatim codes requiring little interpretation.* The total number of verbatim statements for six stories was 262. These 262 were then distilled into 18 general categories by two independent judges and the author. The 18 categories were derived where possible from the attribution literature itself (e.g., effort, luck, task difficulty, help from others, mood) and from the responses themselves (e.g., motivation, time spent, age, facilities). The two independent judges, both psychology students, were presented with category descriptions (e.g., Task Difficulty. Here the explanation for the success or failure is placed outside the individual and located in the task he or she has been asked to perform. For example failure may be explained by the difficulty of the task and success by the ease). The judges were told nothing of the study's rationale, nor the ages of the children whose responses they coded. They knew only the stories which had resulted in the responses they were classifying, that the number of statements falling into each category would probably not be equal, and that statements over which there was uncertainty should be left unclassified. There was majority agreement (i.e., two or three person agreement) on 89 per cent of the statements. Light's (1971) coefficient of three-judge agreement was 0·75 ($P < 0{\cdot}001$). The reliability of coding of the original verbatim statements into the extensive coding system some two years earlier was checked by selecting two transcripts from each age group, one male and one female. A total of 235 recodes was made. There was complete agreement on 95 per cent of the codings.

---

* The stories were designed through a pilot study in another primary school where a small number of children and teachers were interviewed and classrooms observed. In order to ensure that the stories were equally recognisable as achievement events for all age groups the precise content of the achievement event in stories 4 and 5 was changed for the age group 4 children. In story 4 age group 4 children were told that one child was in a higher English 'set' than another child; and in story 5 that one child was in a higher maths 'set' than another. In story 4 also, children in age groups 2 and 3 were asked about their 'maths work' and 'textbooks'.

† It was found in the pilot study that children frequently responded to some of the causal questions with a description of the outcome, e.g., "What is the reason why X is working well?" "He gets all his sums right." This kind of response occurred frequently among the youngest age group. In order to give the child the maximum opportunity to produce an 'explanation' rather than a 'description' a probe question of the form "What is the reason why he got all his sums right?" was asked alongside the other probe of "Can you think of any other reasons?"

RESULTS

The results are presented in three sections. First, the different categories of explanation are described briefly. Second, the frequencies of use of different categories are presented. Third, the effect of age on the use of the categories is analysed through profile and other multivariate analyses.

The categories

The 18 categories into which children's explanations for academic success and failure were classified and an example of each are listed below. A fuller description of the categories is offered in Appendix 1.

1. Performance Ability (e.g., he does nice writing).
2. Specific Competence (e.g., she knows how to do sums).
3. General Competence (e.g., he's intelligent).
4. Previous Achievement (e.g., he did good work in the infants class).
5. Effort and Interest (e.g., she's interested in what she's doing).
6. Motivation (e.g., he wants to get into higher sets).
7. Time Spent (e.g., one started school before the other one).
8. Other Persons (e.g., his parents help him).
9. Task Difficulty (e.g., the work is too easy for students in lower sets).
10. Behaviour (e.g., he mucks about).
11. Age (e.g., she's older than the other one).
12. Facilities (e.g., he had a nice paint brush).
13. Mood and Physical state (e.g., maybe she was in a temper).
14. Domestic Situation (e.g., she's got problems at home).
15. Speed (e.g., she works faster than he does).
16. Sex Stereotype (e.g., girls think that boys are supposed to be good at maths).
17. Personality (e.g., she's a nice girl).
18. Chance (e.g., there is a low chance of being moved up from a low set and a high chance of being moved down from a high set).

---

* A response to story 4 of the kind "well one of them tried harder and might have been clever and the other's mum didn't help him" would have been broken up into three verbatim statements and coded under 'tried harder', 'cleverer' and 'mum didn't help'.

Frequency of use of the categories

Children were classified as having used a category if they used it at least once in response to a particular story. Table 2 presents the frequency of use of the different categories.

TABLE 2

FREQUENCY OF USE OF ATTRIBUTIONAL CATEGORIES BY STORY

	Story						Average rank order
	1	2	3	4	5	6	
Performance Ability	56	22	25	59	36	27	2
Specific Competence	19	8	6	14	17	24	7
General Competence	15	2	9	17	13	50	6
Previous Achievement	1	0	0	9	2	0	12
Effort and Interest	55	29	36	31	33	51	1
Motivation	5	1	3	4	6	7	9
Time Spent	0	3	4	47	37	9	4
Others	8	4	6	15	17	4	8
Task Difficulty	0	0	0	3	10	2	10
Behaviour	22	19	22	11	17	9	4
Age	1	0	1	7	1	3	11
Facilities	0	0	0	0	0	6	13
Mood and Physical State	0	2	0	0	0	0	16
Domestic Situation	0	0	0	1	0	0	17
Speed	4	7	0	40	45	12	3
Sex Stereotype	1	1	2	2	1	0	14
Personality	1	1	0	0	0	1	15
Chance	0	0	0	0	1	0	17

N.B. Children were classified as having used a category if they used it at least once in response to that story. Whether one averages the absolute frequencies or computes the numbers who use the category at least once in response to any story, the rank order remains almost identical.

Table 2 indicates that certain attributional categories occur much more frequently than others in a free-response situation. Effort and Interest and Performance Ability are the most frequently used attributions. The three different ability categories all occur frequently. Somewhat surprisingly perhaps the attribution Behaviour (teacher sanctioned behaviour) is also used very frequently along with Speed of working and Time Spent on the work. Along with Specific Competence, General Competence and Other Persons these are the most frequent attributions. Of much lesser importance — in a free response context — are Task Difficulty, Chance and Mood.

The use of categories as a function of age

Profile analysis (Morrison, 1976) was used to examine whether age had an effect on the use of different attribution categories. The advantage of profile analysis is that it controls for fluency effects — an obvious problem when the free responses of children of different ages are analysed quantitatively. Profile analysis enables one to examine whether age affects all categories equally leaving the overall shape of the category profile unchanged, or whether age affects the shape of the profile. The disadvantage of profile analysis is that the test for independence of age group profiles involves a comparison of the matrix of between-age variances with the matrix of within-age or error variances. The problem arises when the determinant of the within-age matrix is zero, the likelihood of which increases when cell frequencies are zero. Since some categories were never used by some age groups in some stories, and cell frequencies were zero, the profile analysis was applied to the nine most frequently occurring categories — Performance Ability, Specific Competence, General Competence, Effort and Interest, Behaviour, Others, Motivation, Speed, Time Spent. The six stories were analysed separately. Table 3 presents the results of the Profile analyses.

TABLE 3

EFFECTS OF AGE ON THE SHAPE OF ATTRIBUTIONAL CATEGORY PROFILE

	N	Pillais test value	Approx. F value	Hypoth. df	Error df	Significance of F
Story 1	116	0·53	2·84	24	321	0·00002
Story 2	80	0·68	2·58	24	213	0·00015
Story 3	83	0·85	3·64	24	222	0·00001
Story 4	146	0·72	5·39	24	411	0·00001
Story 5	148	0·53	3·73	24	417	0·00001
Story 6	146	0·54	3·76	24	411	0·00001

N.B. Since it was possible in Stories 1-4 for subjects to fail to identify other children who were working well, or working better than others, only those subjects asked the follow-up causal questions were included in the analysis. This explains the varying sample sizes.

An additional multivariate analysis was performed on the pooled data and restricted to children who had been asked the key causal questions in all stories $(N = 70)$. This restricted sample was unbiased with respect to the age, sex and ability characteristics of the main sample (Little, 1982). There was a main effect for age on the response vector. $F(27, 161) = 6{\cdot}01$, $P < 0{\cdot}001$. Repeated measures were used to examine the effect of age on each of the response categories. The univariate F tests are shown below in Table 4.

There are strong age effects for Performance Ability, Specific Competence, General Competence, Effort and Interest, Others and Time Spent. The means indicate that younger children are more likely than older children to use Performance Ability and that older children are more likely to use Specific Competence, General Competence, Effort and Interest, and Others. The relationship between age and group means for Time Spent is curvilinear. Age groups 2 and 3 have higher means than age groups 1 and 4.

TABLE 4

UNIVARIATE F TESTS FOR AGE AND CATEGORY

Attribution category	Univariate F	P<	Direction of means
Performance Ability	7·29	0·0003	negative function of age
Specific Competence	4·09	0·01	positive function of age
General Competence	10·80	0·0001	positive function of age
Effort and Interest	25·58	0·0001	positive function of age
Behaviour	1·69	0·18	
Others	4·37	0·007	positive function of age
Motivation	0·87	0·46	
Speed	1·53	0·22	
Time Spent	6·42	0·008	curvilinear function of age

DISCUSSION

The discussion is organised around the range and frequency of use of categories on the one hand and on age changes in the pattern of use on the other.

The uses of different types of attribution

The overall picture of the types of explanation used freely by schoolchildren lends considerable support to Heider's early (1958) and Weiner's more recent (1971, 1979) statements of the centrality of ability and effort attributions. But at least three qualifications need to be made — the first on different types of ability attribution, the second on the potential importance of Behaviour as an explanation, and the third on the infrequent mention and different meanings attached to luck explanations.

First, an ability attribution can mean several different things, each with potentially different implications for the attribution process. The first type of ability attribution — Performance Ability focuses on the achievement outcome. The child describes observable outcomes and processes of the achievement situation. The second type — Specific Competence — combines a dispositional term like 'know' or 'can' with an achievement act. The third type — General Competence — employs only a dispositional characteristic with no reference to action.

Second, the attribution Behaviour (e.g., he's naughty, she mucks around, he talks during lessons) was used frequently by children to explain classroom achievement events. In none of the free response surveys of adult or fifth grade student attributions (quoted in Weiner, 1979, and Frieze, 1980) has Behaviour been mentioned as an important and significant attribution for schoolchildren. Though such an attribution may be unimportant in the mature, adult and socialised or middle-class classrooms from which subjects for most attribution studies are drawn it appears to be of considerable importance to the children of this author's sample. Classrooms for these children are not only places where one learns 'knowledge'. They are also microcosms of wider society where one learns to 'Behave'. The social control function and the knowledge function interact, and children, at least, perceive the connection.

Third, task difficulty and luck are, for Weiner (1979) very important attributions for success and failure. These types of response feature little in the responses of 5- to 14-year-olds to the questions posed in this research, chance (luck) occurring only once as an explanation throughout the entire exercise. This finding of luck is supported by recent evidence from elsewhere. Fyans and Maehr (1979) noted the infrequent use of luck attributions among American twelfth grade students. For adults, too, the role of luck and chance has also been questioned (Frieze, 1976). A further problem associated with an attribution to luck concerns its meaning.

Because luck was so infrequently used in free-response attributions children in age groups 2, 3 and 4 were asked directly whether luck was a word that could be used when talking about schoolwork. If they said "yes" they were then asked how it could be used. Older children were more likely than younger children to deny that luck could be applied to schoolwork ($b=0{\cdot}21$ $P<0{\cdot}01$). However, among those who did acknowledge its use ($N=60$) there was considerable variety of meaning. Some children (10 per cent) were unable to explain how it could be used. The majority (55 per cent) used the word in an evaluative sense ("he was lucky if he got a star"). A smaller percentage (35 per cent) used the word in a chance sense ("it's like when you just pick the right answer from five possible answers"). This last meaning, the chance meaning, is presumably what attribution theorists have in mind when they classify luck as an attribution which is both external and uncontrollable. Younger children were more likely than older children to use luck in its evaluative sense, while older children were more likely than younger children to use it in its chance sense ($b=-0{\cdot}32$ $P<0{\cdot}01$).

Age changes in the pattern of use

Before the psychological significance of the age trends reported earlier can be discussed, comments should be made on factors besides age which could have produced some of the results.

Story effects

The first obvious point concerns the effect of type of story. The dimensions of the story information (e.g., real *vs.* hypothetical others; success *vs.* failure; relative success *vs.* relative failure; specific *vs.* general achievement events, consistency) will determine specific types of attribution. Others have pointed to some of these effects (Miller and Ross, 1975; Frieze, 1976; Shultz and Butkovsky, 1977; Medway, 1979). Indeed the multivariate analysis reported earlier indicated a main effect of story on the response vector (F (36, 916) = 6·23, P < 0·0001) and a significant age × story interacton (F (108, 1788) = 2·35, P < 0·0001).

Age group differences or different age group environments?

It could also be argued that the age effects result not so much from the fact that children of different ages construe the same world differently but that the classroom environments in which children find themselves are different for the different age groups. To what extent are children being asked to explain the same reality? It was possible to explore this possibility in three ways. First, are children of different ages equally accurate in their perception of academic success and failure? Second, are the 'objective' determinants of 5-year-olds' academic success and failure the same as 14-year-olds'? Third, do teachers explain 5-year-olds' academic success and failure in the same way as they would explain 14-year-olds'?

The first question was explored with data available for age groups 1 and 3. Children identified by subjects as 'working well' were compared with their scores on the Otis-Lennon test of ability. Children in both age groups displayed high and remarkably similar levels of accuracy in their perception of another's success.

The second possible environmental difference — that the 'objective' determinants of achievement differ for the different age groups — is even more interesting. It could be argued, for example, that since all 5-year-olds appear to try equally hard in their first year of schooling then effort is not available to them as an explanation for differential achievement. This could explain why they use effort so little in their explanations. Conversely, it could be argued that many 14-year-olds are disillusioned with their schooling and that there is considerable variation between individuals in the amount of effort they invest, and that effort is therefore more readily available as an explanation. It was possible to examine partially the 'objective' determinants of achievement among the youngest and oldest age group through effort and achievement ratings from school files and ability scores from the Otis-Lennon test. Among the youngest children effort was a much more important determinant of achievement than was ability and was more powerful than the effect of effort among the oldest children. Among the oldest children ability was more powerful than effort and was more powerful than the effect of ability among the youngest children. While confirming the point that the 'objective realities' of the causes of achievement are different for younger and older groups the data are somewhat contrary to those expected. The subjective explanations of 14-year-olds emphasised the importance of effort. Objectively, effort was not very important. Conversely the subjective explanations of 5-year-olds included effort but not very frequently. Objectively the impact of effort on achievement was very strong.

A third possible environmental difference on which data were available was the attributions made by the subjects' teachers. The six teachers responsible for the sample children in age groups 1, 2 and 3 were asked to rank in order the importance of home background, effort, class behaviour, native intelligence, teacher attention and luck for above-average and below-average performance of the pupils in their class. Effort was equally important for teachers of all three age groups for both success and failure. Moreover, where there were differences between teachers of different age groups and between teachers of the same age group there was no relationship between the teachers' attributional preferences and children's propensity to use the attribution.

These three sets of evidence certainly do not deny the importance of the social environment in which children explain academic achievement. Indeed it would be implausible to suggest that children could explain achievement without a social environment made up of consensual definitions of achievement, achievement values, 'objective' determinants of achievement and teacher and parent beliefs about the importance and determinants of academic achievement. What this analysis does suggest is that children's explanations do not simply reflect an objective classroom reality. Rather, they are determined through a complex interaction between objective and subjective reality.

Consistency with other developmental characteristics of social cognition

Five changes in that subjective reality — the decreasing importance of Performance Ability and the increasing importance of Specific Competence, General Competence, Effort and Interest, Others — are consistent with other developmental characteristics of social cognition. The characteristic of the Performance Ability category is its descriptive rather than explanatory nature. The child appears to be preoccupied with coming to terms with differential achievement, with describing and redescribing it before seeing it as something in need of explanation. Jill (5.11) demonstrates quite clearly her determination to 'explain' an event through continued description. For her the reason why one child paints a nice picture while another child paints not such a nice one is

— *because it's done very nice*
— *What about the other one?*
— *The other one means that it's horrible.*
— *Yes, but do you know the reason why one has done it nice and one has done it horrible?*
— *Yes, the nice one is much nicer and the horrible one, it's much horribler.*

The performance ability attribution reflects Heider's (1958) notion of different 'levels of attribution'.

> The total relevance of the event so encompasses a manifold of attributions of varying depth (p. 255) . . . the search for relatively enduring aspects of our world . . . may carry us quite far from the immediate facts or they may end hardly a step from them.

This preoccupation of young children with describing an event in terms of another observable event is consistent with several other developmental characteristics of children. During the stage of 'pre-operations' Piaget and Inhelder maintain that the child's cognitions are dependent on the observable, the concrete and the immediate (Inhelder and Piaget, 1958) while the work on moral judgment and achievement evaluation emphasises the importance for the young child of outcome as a basis for evaluation (Piaget, 1932; Weiner and Peter, 1973; Leichmann, 1976). In their study of children's understanding of social institutions Furth *et al.* (1976) found that many children around the age of 6 years displayed what they called non-inference or non-understanding. The child confines reasoning to observed events. Observed events and observable characteristics are also central to the analysis of data in the person perception studies. Young children's descriptions of other people focus on the static and the concrete and rarely move beyond to the abstract and the psychological (Livesley and Bromley, 1973). Wood (1978) describes changes in the child's understanding of other people's motives for behaviour. The remarks of the youngest children were characterised by descriptions of immediately observable behaviour while older children moved beyond description to the attribution of personality type.

A performance ability attribution represents the first step in a search for enduring aspects of the world of achievement. First, the 'fact' of achievement must be grasped and described before the search for dispositional properties of the person can be sought. The search for two dispositional properties — specific and general competence is highlighted next.

In contrast to the negative age trend for performance ability the two categories of specific and general competence show positive trends. The categories of specific and general competence represent a shift of emphasis away from observable action and towards the person. This shift away from observable action to explanations in terms of the inner properties and competencies of the individual is consistent with Piaget's characterisation of the movement from the pre-operational to the concrete operational to the formal operational stages of cognitive development. In the pre-operational stage, thought is intimately tied up with action. In the concrete operational stage the child can think beyond the immediately observable but the abstractions and operations which are performed are still very dependent on the observable. In formal operations however the child can operate on abstract thoughts. The concrete representations of those thoughts do not have to be present for those thoughts to continue. The child can think about thinking. In performance ability the explanation is tied to the immediate and the observable.

she is doing her work nicely.

In specific competence there is some abstraction to the inner properties and predispositions of the person but the abstraction is still linked to a specific action.

he's good at doing sums.

In the general competence statements, however, the link with action is broken and a statement about the inner properties of the person provides an explanation of behaviour. There is no explicit reference to action or behaviour.

she's intelligent.

The Effort and Interest Category yielded the strongest univariate effect (Table 4). The increase in use with age is quite remarkable. The belief that success and failure are due to effort and lack of effort is well established by the age of 11. This trend parallels the other well known trends in the use of effort rather than outcome as the basis for reward decisions (Piaget, 1932; Kohlberg, 1963; Weiner and Peter, 1973; Rogers, 1980) but its increasing use with age as an attribution for success and failure has not been discussed before.

Finally, the use of the category Others increases with age. Included here are attributional statements of the kind "his parents help him", "the teacher was strict with him", "his brothers help him at home". In one sense these explanations approximate Piaget's true causal explanations for physical phenomena, the use of external forces to explain physical events. The increasing use of this kind of explanation parallels findings from children's free descriptions of others. Shantz (1975), summarising the main free description studies, observes that children older than 8 make increasing use of interpersonal characteristics, i.e., descriptions of the other in terms of his or her effect on other people.

So far it has been suggested that Weiner's models of attribution require modification when applied to schoolchildren making sense of achievement events in the classroom. First, the types of attribution which children typically make are more extensive than those normally included in attribution experiments, that two of those often included (luck and task difficulty) occur infrequently in children's responses to free response questions and that types of ability attributions need to be distinguished. Second, while recognising the importance of the situational determinants of attributions, it is suggested that the child's subjective reality and objective reality interact to produce different attributions for children at different stages of psychological development.

A consideration of these two points — that the range of attributions used by children in the classroom may be greater than presumed by current models and that the most salient attributions to be included in any model may change over time — also raises a question about the psychological organisation of attributions. Weiner (1979) has presented a taxonomy of attributions which involves more than a listing of types of attribution. It involves the organisation of these attributions along dimensions of locus, stability and controllability. These dimensions are imposed on the attributions by the psychologist. They represent *his* sets of meanings, his way of interpreting the world. But since attribution theory is essentially concerned with understanding the lay person and lay child's understanding should not the child be encouraged to display his or her own organising system? Children may simply not view low ability as uncontrollable and stable. Many children, aware of the rapid development of his or her own abilities through the early school years, may perceive ability to be unstable and controllable through effort. From whose point of view is the attribution 'clever' both stable and uncontrollable? Weiner's (1979) classification of 'clever' as internal, stable and uncontrollable receives substantial support from dominant psychological traditions which conceive of ability as largely (though not wholly) pre-programmed genetically and giving rise to stable and uncontrollable individual differences. But at what point should the attribution theorists 'etic' be permitted to interfere with the child's 'emic'? (Berry, 1969). Future research on attribution in children should seek to establish the validity of the second order dimensions currently used to organise attributions.

CONCLUSIONS

This study has examined the adequacy of current models of attribution for British schoolchildren. Using a free response interview a range of attributions emerged. It was suggested that the models designed originally to encapsulate adults' attributions be refined. Specifically, ability attributions should be differentiated into more than one type, the attribution luck is not very important for schoolchildren and means different things for different children, children frequently allude to good and bad behaviour as determinants of academic success and failure. It was also suggested that taxonomical classifications of attributions should make clear whose taxonomy is being used — the psychologist's or the child's? If one is interested in the implicatons of attributions for the child's behaviour then the child's own organising system may be of much greater importance than the psychologist's. The psychologist's task is to elicit from the child his or her own dimensions.

Developmental trends in the use of different attributions were apparent. The pattern of use of different attributions by the different age groups was examined through profile analysis. The pattern was different for the different age groups. The use of the performance ability category decreased with age while the Specific and General competence categories, Effort and Interest and Others increased with age. These changes were related to other developmental changes in social cognition (Inhelder and Piaget, 1958; Livesley and Bromley, 1973; Furth *et al.*, 1976). These developmental trends were discussed alongside the possibility that the objective stimuli to which children of different ages are exposed are sufficiently different to explain the results. Although there were differences in the environment these were sometimes contrary to what one might expect It is concluded that children's explanations do not simply reflect objective reality. The interaction between objective and subjective reality is complex.

REFERENCES

BERRY, J W. (1969) On cross-cultural comparability. *Int J Psychol*, 4, (2), 119-28

FINCHAM, F. D (1982). Developmental dimensions of attribution theory. In JASPARS, J., HEWSTONE, F , and FINCHAM, F (Eds.), *Attribution Theory and Research*. London· Academic Press

FRIEZE, I (1976). The role of information-processing in making causal attributions for success and failure In CARROLL, J. S , and PAYNE, J W (Eds) *Cognition and Social Behaviour*. Hillsdale, N J. Erlbaum.

Frieze, I. R (1980). Beliefs about success and failure in the classroom. In McMillan, J. H. (Ed.), *The Social Psychology of School Learning*. New York· Academic Press.

Furth, H. G , Baur, M., and Smith, J. E. (1976). Children's conceptions of social institutions: a Piagetian framework *Human Dev.*, 19, 351-374.

Fyans, L J., and Maehr, M L. (1979). Attributional style, task selection and achievement. *J educ. Psychol.*, 71, 499-507.

Heider, F (1958) *The Psychology of Interpersonal Relations*. London: Wiley.

Inhelder, B., and Piaget, J. (1958). *The Growth of Logical Thinking from Childhood to Adolescence* New York. Basic Books.

Kelley, H. H , and Michela, (1980). Attribution theory and research. In Rosenzweig, M. R., and Porter, L. M. (Eds.), *Ann. Rev. of Psychol.*, 31.

Kohlberg, L (1963). The development of children's orientations to a moral order· 1. Sequence in the development of moral thought. *Vita Humana*, 6, 11-33.

Leichmann, G. A (1976) The effect of age and educational environment on the development of achievement evaluations and moral judgements. Unpublished Ph.D. Thesis, London School of Economics.

Light, R J (1971) Measures of response agreement for qualitative data· some generalisations and alternatives *Psychol. Bull.*, 76, 365-377

Little, A. W (1982). The development of the child's understanding of the causes of academic success and failure Studies of English and Sri Lankan schoolchildren. Unpublished D Phil Thesis, University of Sussex.

Livesley, W. J., and Bromley, D B (1973) *Person Perception in Childhood and Adolescence*. London: Wiley.

Medway, F. J. (1979) Developmental aspects of self and other causal attributions. *Psychol Rep* , 45, 155-159.

Miller, D. T , and Ross, M. (1975). Self-serving biases in the attribution of causality. fact or fiction? *Psychol Bull.*, 82, 213-225.

Morrison, D. F. (1976) *Multivariate Statistical Methods*. London. McGraw-Hill.

Piaget, J (1932). *The Moral Judgement of the Child*. London: Routledge, Kegan and Paul.

Rogers, C. G (1980). The development of sex differences in evaluations of others' successes and failures *Br. J. educ. Psychol.*, 50, 243-252

Shantz, C U. (1975) The development of social cognition In Hetherington, E. (Ed), *Review of Child Development Research — 5* Chicago University of Chicago Press.

Shultz, T. R., and Butkovsky, I (1977). Young children's use of the scheme for multiple sufficient causes in the attribution of real and hypothetical behaviour *Child Dev.*, 48, 464-469.

Weiner, B., Frieze, I H., Kukla, A., Reed, L., Rest, S. and, Rosenbaum, R. M. (1971). *Perceiving the Causes of Success and Failure* General Learning Press, Morristown, NJ General Learning Press.

Weiner, B., and Peter, N. (1973). A cognitive-developmental analysis of achievement and moral judgements *Dev Psychol.*, 9, 290-390.

Weiner, B , and Kun, A. (1976). The development of causal attributions and the growth of achievement and social motivation. In Feldman, S , and Bush, D. (Eds.), *Cognitive Development and Social Development*. Hillsdale, NJ · Erlbaum.

Weiner, B (1979) A theory of motivation for some classroom experiences. *J educ. Psychol.*, 71, 3-25.

Weiner, B., Kun, A., and Benesh-Weiner, M. (1980) The development of mastery, emotions and morality from an attributional perspective In Collins, W. A (Ed), *Development of Cognition, Affect and Social Relations*. The Minnesota Symposium on Child Psychology, 13 Hillsdale, NJ : Erlbaum

Wood, M. (1978). Children's developing understanding of other people's motives for behaviour *Dev Psychol.*, 14, 561-562.

(*Manuscript received 29th February*, 1984)

APPENDIX

ATTRIBUTIONAL CATEGORY SYSTEM

Eighteen general categories of explanation were used to classify the free responses of children. These are listed without priority below.

The first four categories are different types of ability statements. From the very beginning of this research difficulties were encountered in the classification of ability statements. As far as one can discern most researchers into developmental attribution processes include statements as diverse as "he does his maths work well", "he knows how to do maths" and "he's intelligent" in the same ability attribution This global classification is considered unsatisfactory since these statements seem to vary on a number of dimensions, e.g., their specificity of ability, their time reference, their reference to action and their reference to adjectival properties of the person. The first four categories listed below are different types of ability statements.

P 3889

(i) *Performance Ability* These responses describe the observable outcomes and processes of the achievement situation. For example

— *What is the reason why he is working well?*
— *Because he does nice writing*

and

— *What is the reason why Darren is working better than Trevor?*
— *It's because he never gets his sums wrong.*

To the adult eye these responses do not seem to be 'real reasons' or explanations. They appear to be redefinitions of the question. They involve describing the achievement event or outcome which the child has been asked to explain. And where the statements do go beyond the outcome they focus instead on the observable process of achieving that outcome, e.g., he writes neatly, he draws carefully. Included in this category also are tautological statements. For example

— *What is the reason why one child is on reading book six and the other is on reading book five?*
— *Because one is on a higher book than the other one.*

and

— *What is the reason why one child has finished his number book and the other has half finished?*
— *Because one child has finished more pages.*

They appear to be tautologies to the adult because the responses do not go beyond the information presented in the story. Implicit in the information that one child is on reading book six and one is on reading book five is (for an adult) the notion that one book is of a higher grade; that to be on book six means that book five has already been read. But to young children, who are not terribly good at 'reading between the lines', and inferring, these explanations have certainly gone beyond the given information. After all it is not as though the child responded that the reason why one child was on reading book six and the other on five was because one was on six and one was on five. Rather, the recognition that one is on a higher book has involved going beyond the given information. The child moves only as far as the outcome itself however. The outcome is redescribed.

(ii) *Specific Competence Ability* These statements *combine* a 'dispositional' term like 'know' or 'can' with an achievement action which is part of the final outcome. For example

— *he knows how to do sums*
— *he can write quickly*
— *he's good at writing*

This last example can be distinguished from 'he writes well'. Here there is no use of a dispositional term and the words used describe only what the child does "He writes well" would be classified under Performance Ability and not under Specific Competence. The dispositional term implies a certain competence on the part of the person and the action term ensures that the competence is not generalised to all achievement activity. It is specific to a certain set of activities. This dissociation of the person from action and the link between them is neatly expressed by the child who explained another child's lack of achievement by the fact that "he takes time to cling on to the work".

(iii) *General Competence Ability* If specific competence links a person's characteristics with his or her actions then general competence frees the person from those specific actions. Behaviour is explained not by a description of that behaviour nor by a combination of predisposition and behaviour but solely by a personal characteristic, e.g., "he's brainy"; "she's clever"; "she's intelligent", "he's thick". There is no reference to what the person can do or does do.

(iv) *Previous Achievement* A number of children made explicit reference to past achievement in their explanations of present achievement, e.g., "it depends on how well you did in the primary school"; "they did good work in the infants class"; "she did more at the old school" Some of these statements could arguably be placed within the category called Performance Ability in that they refer to actions of the child. However, the reference to historic time is so explicit that the statements warrant separate classification.

(v) *Effort and Interest* Grouped together in this category are statements like "he tries harder"; "she gets on with it"; "she's interested in what she's doing"; and "he likes that subject".

All the statements imply a conscious involvement (or lack of) in the achievement event, an involvement which is not immediately observable but rather is inferred from observable behaviour.

(vi) *Motivation* These are statements of rough equivalence to Piaget's causal category of Teleological explanation. One does something in order to achieve something else in the future, e.g., "they think that if they work hard they will get a good job in the future"; "she wants to get into higher sets"; "she wants to get a star". The goal to be achieved in the future explains the event in the present.

(vii) *Time Spent* This category refers either to the time spent in school, the time lost through illness or the time spent by one child on a particular activity in the class, e.g., one started school before the other one; one's older than the other one; one started on a lower book; one child came to that class later.

(viii) *Other Persons* These explanations involve the effects of the actions or attitudes of other persons on the child, e.g., his parents help him; the teacher was strict with him; the teacher helped her more; his brothers help him at home; he had a better teacher in the primary school.

(ix) *Task Difficulty* There were a number of variations on this theme, e.g., the work is too difficult; the work is too easy for students in lower sets; one book was easier than the other one. The explanation for the low or high achievement is to be found in the task set rather than in the child's effort or ability to handle it. At times this sort of explanation appears to be illogical to the adult eye particularly when the two children in a story are performing identical tasks.

(x) *Behaviour* These statements cover those actions sanctioned by the teacher, actions which are often referred to as 'bad behaviour' or behaviour 'problems' e.g., "he mucks about"; "he's naughty"; "he talks during lessons". Also included in this category are statements like "he copies".

(xi) *Age* Some children explained success and failure by referring to the age of the child, e.g., he's older than the other one; he's the youngest. To an adult these statements might logically be reduced to category (vii), Time spent on the work, but this is an invalid classification for the 6-year-old child for whom age is often a non-reducible explanation, an explanation unto itself.

(xii) *Facilities* The instruments and facilities available to children e.g., blunt pencils; nice paint brushes, were also cited as contributing to high and low achievement.

(xiii) *Mood and Physical State* This category includes statements like "he's excited"; "one's happy, one's angry"; "maybe she's in a temper"; "she feels sick". Such statements carry with them the connotation that the states of temper, anger and happiness are not permanent.

(xiv) *Domestic Situation* There were a few references to one's situation at home. "She's got problems at home" was sometimes mentioned. Further probing did not reveal the nature of these problems. However, other statements of the kind "her parents help her with her schoolwork" or "her parents do not help her" were not classified under this heading. Instead the interactional dimension of *others' influence* took precedence and these statements were classified instead under category (viii).

(xv) *Speed of Working* In these statements the child uses an adverb or adjective to express the speed with which a child works, e.g., "he works quickly"; "she works faster than he does"; "he's a slow worker". This category also includes references to high speed which lead to low achievement, e.g., "he hurries his work".

(xvi) *Sex Stereotype* Some children used statements of the sort "girls tend to be in higher sets"; "boys tend to be in higher maths sets"; "girls think that boys are supposed to be good at maths"; the category includes not only the single stereotype "girls tend to . . . and boys tend to . . ." but also the reflexive stereotype "girls think that boys tend to".

(xvii) *Personality* In this category are statements like "she's quiet"; "she's a nicer girl".

(xviii) *Chance* There was just one reference to chance throughout the entire interviewing on these particular questions. One child, in response to a question about the reasons why one child was in a lower group for mathematics work than another replied that there is a low chance of being moved up from a low set and a high chance of being moved down from a high set.

Br. J. educ. Psychol., 55, 24-27, 1985

RIGID PERSONALITY AND RIGID THINKING

BY P. KLINE AND C. COOPER
(Department of Psychology, University of Exeter)

SUMMARY. Tests of rigid thinking and rigid personality were administered to a sample of 173 undergraduates. Correlations, factor analysis and canonical analysis all indicated that there was little relationship between these two sets of variables.

INTRODUCTION

IN the field of abilities, rigid thinking is a well-established factor, albeit under a variety of names. Guilford and his colleagues (e.g., Guilford, 1967) refer to its negative pole as divergent thinking which is, of course, contrasted with convergent thinking. Despite the problems with Guilford's three faceted model of intellect on account of the difficulties with his Procrustes rotations (Horn and Knapp, 1973), there is little doubt that divergent thinking abilities, flexibility, fluency and originality, do load a factor. Cattell (e.g., 1971) has labelled it the retrieval factor. Kline and Cooper (1984b) found that this emerged with clarity in their study of the Comprehensive Ability Battery (Hakstian and Cattell, 1976). Horn (1976) regards it as a marker factor in the field of abilities as does Carroll (1983).

Psychologists concerned with the development of creative thinking in its widest sense, for example Wallach and Kogan (1965), De Bono (1967), and Hudson (1969), despite the lack of clear evidence have all assumed that divergent thinking abilities are important in real life creative activity. More recently Dixon (1976) in his study of military incompetence has emphasised the importance of flexible thinking in the good conduct of war.

Rigidity, however, is also a term that is used together with its synonyms in the description of personality. Rigidity, authoritarianism, dogmatism, obsessionality and inflexibility (not to say anal character) are well known constructs, each with their own scales. Dixon (1976) argued that all these were essentially measuring the same variable, a claim supported by Kline and Cooper (1984a) in their study of the authoritarian personality. Dixon also argued that these authoritarian traits contributed to military incompetence, at least at a high level of command. Indeed, it is implied there, as it is by Hudson (1969), that rigid thinking and personality are related. Certainly in the original study of the authoritarian personality by Adorno and colleagues (1950) and in the study of dogmatism by Rokeach (1960) there is the strong implicaton that rigid, inflexible persons are not creative.

In the present paper we therefore examine the relationship between rigid personality and rigid thinking.

Subjects

173 undergraduates from all faculties of the university participated in this study. Approximately half were male, half female. Their mean age was 20·4 years and the standard deviation was 1·4 years. All were volunteers, taking part in a larger study involving 12 hours of testing. The reward to subjects was individual feedback concerning their test results.

Tests

As indicated above, a total of 12 hours testing was carried out, the tests being divided into 4 three-hour sessions. In the present study the test variables were:

Personality

(*a*) Ai3Q (Kline, 1971): this is a test of anal character or obsessional traits.

(*b*) The Wilson-Patterson Conservatism Scale: this scale is clearly related to conservatism and measures inter alia ethnocentrism, anti-hedonism, religiosity and militarism (Wilson, 1975). These four scales were used in this study, the others being unsuitable because they shared items.

(*c*) Balanced Dogmatism (Ray, 1974): this scale captures the essence of the original authoritarian scale but overcomes many of its psychometric problems.

(*d*) Kohn's balanced F scale (Kohn, 1972): a further psychometrically improved F scale.

Abilities

Primary abilities were measured by the Comprehensive Ability Battery (Hakstian and Cattell, 1976). Subjects took 17 of these tests — however only the six which were clearly related to flexibility were included in the present analysis. Spontaneous flexibility, ideational fluency, word fluency and originality were clearly related to flexible thinking since they were designed to load on Cattell's retrieval factor. Speed of closure and flexibility of closure were also investigated in this study.

Statistical analysis

For the purposes of our study, correlations between the seven measures of rigid personality and rigid thinking were computed. This was probably sufficient to establish the degree of overlap between ability and personality but as a further check a Direct Oblimin (Jennrich and Sampson, 1966) simple structure oblique factor analysis was performed, together with a canonical correlation analysis between the two sets of variables, this latter technique serving to maximise the degree of overlap between the two sets of variables. Non-conservative significance levels of 0·05 were adopted when comparing correlations; although it is then quite probable that "false-positives" would be found, a failure to reveal any substantive overlap between the personality and ability variables could not be ascribed to the use of an unduly conservative significance level.

RESULTS AND DISCUSSION

Ai3Q, Militarism, Anti-Hedonism and Religiosity had no significant correlations with any of the creativity variables. Ethnocentrism had a negative correlation with word fluency of $-0{\cdot}19$ ($P<0{\cdot}01$). Kohn's F scale correlated 0·125 with flexibility of closure ($P<0{\cdot}05$) and Balanced Dogmatism correlated $-0{\cdot}16$ ($P<0{\cdot}05$) with word fluency and $-0{\cdot}19$ ($P<0{\cdot}01$) with originality. None of the remaining 38 correlations was significant.

Although all the significant correlations — except that between the F-scale and flexibility of closure — were in the right direction, two points stand out. Most of the personality variables did not correlate significantly with the creativity variables, as would be expected from the descriptions of rigid thinking and rigid personality. Furthermore the statistically significant correlations were small, so small indeed that their psychological significance is negligible. Thus the correlational analysis indicates that rigid thinking and rigid personality are not closely related.

None of the canonical variates proved significant at the 0·05 level: the first variate showed a canonical correlation of only 0·33 (chi-square$=32{\cdot}3$, df$=36$, $P=0{\cdot}64$). Furthermore, the *signs* of the loadings on the first canonical variate were not consistent. Some flexibility measures showed positive loadings whilst others were negative. This analysis thus reveals little overlap between the personality and ability variables.

In a previous study of Cattell's Objective Analytic Battery (Kline and Cooper, 1984c) an obsessional trait factor loading on all the authoritarian scales used in this study was obtained. No creativity tests loaded on this factor. This factor analytical study supports the more detailed analyses which we have discussed above.

The failure to obtain the expected correlations could be due to an artefact, always a possibility in the study of abilities with student samples, namely attenuation of correlations due to restriction of range on ability variables. However, with this sample this cannot be the explanation since in previous studies both the ability variables and the authoritarian tests loaded on their expected factors. For example, a factor of authoritarianism was found with Ai3Q 0·6, the F-scale 0·84, and G of the 16PF 0·67; in the case of the ability variables there was a clear retrieval factor loading on spontaneous flexibility 0·516, ideational fluency 0·987, word fluency 0·398, and originality 0·434. Factor B (intelligence) did not load on this factor (Kline and Cooper, 1984a, 1984b).

CONCLUSIONS

These results show clearly that rigid thinking, as measured by typical divergent thinking tests, is not related to rigid personality, as defined by measures of obsessional traits, dogmatism and authoritarianism.

This finding has important implications for the training and selection of flexible thinkers. For example the fact that, as Dixon (1976) has shown, authoritarian personalities tend to be attracted into the army by no means precludes the development of flexibly thinking army officers. Research into this problem is now under way. Similarly the fostering of creative gifts in schoolchildren would not appear to be dependent on personality variables, thus allowing teachers to concentrate upon training procedures.

REFERENCES

Adorno, T. W., Frenkel-Brunswick, E., Levenson, D. J., and Sanford, R. N. (1950). *The Authoritarian Personality*. New York: Harper.

Carroll, J. B (1983). Individual differences in cognitive abilities. In Irvine, S. H., and Berry, J. W. (Eds.), *Human Assessment and Cultural Factors*. New York: Plenum Press.

Cattell, R. B. (1971) *Abilities. Their Structure Growth and Action*. New York· Houghton Mifflin.

De Bono, E. (1967). *The Use of Lateral Thinking*. London: Cape.

Dixon, N. F (1976). *On the Psychology of Military Incompetence*. London: Jonathan Cape.

Guilford, J. P. (1967). *The Nature of Human Intelligence*. New York: McGraw-Hill.

Horn, J. L. (1976). Human abilities: a review of research and theory in the early seventies. *Ann. Rev. Psychol.*, 27, 437-485.

Horn, J. L., and Knapp, J. R. (1973) On the subjective character of the empirical base of Guilford's structure of intellect model. *Psychol. Bull.*, 80, 33-43.

Hakstian, A. R., and Cattell, R B. (1976). *The Comprehensive Ability Battery*. Champaign Ill.· Institute for Personality and Ability Testing

Hudson, L (1969). *Contrary Imaginations*. London: Methuen.

Jennrich, R I, and Sampson, P F. (1966). Rotation for simple loadings. *Psychometrika*, 31, 313-323.

Kline, P. (1971). *Ai3Q Test*. Windsor National Foundation for Educational Research.

Kline, P., and Cooper, C. (1984a). A factor analytic study of the authoritarian personality. *Br. J. Psychol.*, 75, 171-176.

Kline, P., and Cooper, C (1984b). The factor structure of the Comprehensive Ability Battery *Br. J. educ. Psychol*, 54, 106-110.

Kline, P., and Cooper, C. (1984c). A construct validity study of the Objective-Analytic Battery. *Person. individ Diff.*, 5, 323-337.

Kohn, P. M. (1972). The Authoritarian-Rebellion Scale. a balanced F scale with left-wing reversals. *Sociometry*, 35, 176-189.

Ray, J. J. (1974) Balanced dogmatism scales. *Aust. J. Psychol.*, 26, 9-14.

ROKEACH, M. (1960). *The Open and Closed Mind.* New York: Basic Books.
WALLACH, M. A., and KOGAN, N. (1965). *Modes of Thinking in Young Children: a Study of the Intelligence-Creativity Distinction.* New York: Holt, Rinehart and Winston.
WILSON, G. D. (1975). *Manual for the Wilson-Patterson Attitude Inventory (WPAI).* Windsor: NFER.

(*Manuscript recieved 7th March,* 1984)

Br. J. educ. Psychol., 55, 28-33, 1985

COGNITIVE SLIPPAGE, TEST ANXIETY, AND RESPONSES IN A GROUP TESTING SITUATION

BY GREGORY C. R. YATES, GLYNIS HANNELL AND R. MARK LIPPETT
(South Australian College of Advanced Education, Magill Campus)

SUMMARY. In this investigation 72 female tertiary students completed questionnaires of their experienced cognitive interference, mind-wandering, enjoyment, and task difficulty following the administration of Set I, and again following Set II, of the Advanced Progressive Matrices. Subjects were informed that Set I was a practice test. The Cognitive Failures Questionnaire (Broadbent *et al.*, 1982) predicted levels of cognitive interference and mind-wandering on Set II. The Test Anxiety Scale (Sarason, 1978) predicted cognitive interference, as well as actual test performance, on Set II. The results were interpreted in line with the view that the CFQ and TAS predict different aspects of the manner in which individuals react in stressful testing situations.

INTRODUCTION

RECENTLY, Broadbent *et al.* (1982) published the Cognitive Failures Questionnaire (CFQ), a self-report device concerned with aspects of cognitive slippage. In using the term cognitive slippage we refer to lapses of attention, processing, memory, perception, and decision-making, in much the same manner as a layman might use concepts such as 'forgetfulness', 'absent-minded', 'inattention', or simply 'lack of attention'. As such, cognitive failure or slippage does not reflect a lack of underlying ability or competency.

In their article, Broadbent *et al.* reported data, from several different samples, indicating that responses on the CFQ (*a*) possess both internal and temporal reliability, (*b*) are uncorrelated with social desirability and lie scale items, (*c*) are uncorrelated with scores on the Progressive Matrices, (*d*) correlate with relevant ratings made by others, (*e*) correlate strongly with other measures of forgetfulness and absentmindedness, and (*f*) correlate at low but significant levels with a locus of control scale, a measure of generalised anxiety, a neuroticism scale, and with various symptoms on a psychiatric questionnaire measure.

The stability of CFQ data, together with their meaningful pattern of intercorrelations with other measures, suggested to Broadbent *et al.* that the CFQ possesses trait (as distinct from state) related properties. In one set of data, based on a sample of student nurses, they found that when the CFQ was administered at six-week intervals it failed to reflect known changes in the nurses' occupational stress levels. However, the questionnaire did predict the manner in which the nurses reacted towards increased stress. Specifically, the CFQ meaningfully predicted psychiatric symptoms upon the Middlesex Hospital Questionnaire (MHQ; Crown and Crisp, 1966) in nurses working in high stress wards, but this result was not found in nurses working in low stress wards. Thus, on the basis of this and other psychiatric data, Broadbent *et al.* suggested that "cognitive failure makes a person vulnerable to showing bad effects of stress, rather than itself resulting from stress" (1982, p. 1).

The present study was designed to investigate further the predictive value of the CFQ as an indicator of psychological vulnerability in a stress-inducing context, i.e., the standard formal testing situation. Testing situations, involving obvious evaluations on the part of authority figures and rigid time constraints, are readily identified as tension-producing experiences which pose particular threats to individuals who exhibit high levels of test anxiety. Highly test anxious (HTA) college students, for example, perform relatively poorly in tests involving experimental tasks, such as anagrams, when tested under highly evaluative conditions (Sarason, 1972, 1978). On the other hand, low test anxious (LTA) students are not particularly disadvantaged when tested under such formal conditions.

The work of Irwin Sarason (1972, 1978), in particular, has uncovered systematic differences in the manner LTA and HTA individuals react towards testing contexts. The reduction in performance of HTA students under external evaluation is associated with a mental set reflecting self-monitoring, negative appraisals, and other self-focused thoughts, ruminations and premonitions, which diverts attention away from task relevant cues and considerations. According to Sarason, "whereas the less test-anxious person plunges into a task when he thinks he is being evaluated, the highly test-anxious individual plunges inward" (1972, p. 393).

As one index of this mental set, Sarason developed the Cognitive Interference Questionnaire (CIQ), upon which subjects are able to report level of mental interference experienced on an immediately preceding task. In an experimental study, Sarason and Stoops (1978) found that HTA students reported elevated levels of cognitive interference under testing conditions involving external evaluation, whereas interference levels in LTA students were significantly lower than the levels for HTA students, and did not vary across evaluation contexts. In other words, cognitive interference tended to be low in LTA individuals irrespective of external evaluation, whereas interference was at a moderately high level in HTA students even under conditions of minimal evaluation, and increased further under conditions of obvious evaluation.

In the present study, female tertiary students completed the CFQ and Sarason's (1978) Text Anxiety Scale (TAS) a week prior to being subjected to the Advanced Progressive Matrices test (Raven, 1962). Within the testing session, subjects were informed that Set I of the matrices was a practice test for Set II. Cognitive interference levels were assessed immediately after administration of Set I, and again after Set II. Besides using the CIQ as the measure of cognitive interference, additional ratings were obtained on (*a*) degree of mind-wandering, (*b*) level of task enjoyment, and (*c*) perception of task difficulty. These latter two ratings served primarily as independent checks on the relative stress potential of the two test conditions, on the assumption that undergoing the full test would prove more stressful than working on the practice test items.

In accord with test anxiety theory (Sarason, 1978), it was anticipated that HTA students, relative to LTA students, would evidence higher CIQ scores across both assessment contexts, Set I and Set II. In accord with the 'vulnerability' hypothesis of Broadbent *et al.* (1982), it was predicted that CFQ scores would predict CIQ levels in the more stressful testing condition, but would be less likely to relate to cognitive interference on the practice items. Of considerable interest in the present study was the nature of the relationship between TAS and CFQ responses. Broadbent *et al.* (1982) had reported a significant correlation ($r = 0{\cdot}31$) between CFQ data and a generalised trait anxiety measure in their nurse sample, and in this and other samples, CFQ data correlated significantly with anxiety symptom scores on the MHQ psychiatric questionnaire. Such data, however, has little bearing on predicting a direct relationship between cognitive failure and test anxiety. Insofar as the TAS and CFQ represent relatively independent dimensions it would be possible to divide a sample of subjects into four classification groups around the median on each questionnaire. In the present study this was found to be a feasible procedure for subsequent analyses.

Thus, in sum, it was anticipated that (*a*) TAS data would predict cognitive interference on both Set I and Set II of the matrices, on the assumption that the TAS identifies individuals who respond quickly and characteristically with elevated tension, (*b*) CFQ would predict cognitive interference in the second, more demanding, assessment, on the assumption that the CFQ identifies individuals who begin to show stress-related symptoms only after some period of time under stressful conditions, and finally (*c*) it was generally predicted that cognitive interference would correlate negatively with actual performance on the matrices test (i.e., Set II).

METHOD

Subjects. The subjects were 72 female students engaged in the second year of an undergraduate diploma of education course. The actual age range was from 18 to 36 years but approximately 80 per cent of the sample were under 21 years.

Questionnaires. The Cognitive Failures Questionnaire was used exactly in the form as published by Broadbent *et al.* (1982). In the present sample the mean was found to be 49·64, with a standard deviation of 10·5 and a median value of 50·5. Individuals scoring 48 or lower were defined as being low in cognitive failure, and individuals scoring 52 or higher were defined as being high in cognitive failure. Two sample items from the CFQ are, "Do you fail to listen to people's names when you are meeting them", and "Do you start doing one thing at home and get distracted into doing something else (unintentionally)". Respondents are required to rate themselves on 25 items using a 5-point scale.

In the present sample, the mean on the Test Anxiety Scale (Sarason, 1978) was found to be 19·67 and a median of 20. Individuals scoring 18 or lower, and 22 or higher, were defined as low, and high, respectively, in test anxiety. The use of these cut-off margins on CFQ and TAS scores resulted in the loss of 10 subjects' data in analyses based on classification groupings. However, correlational analyses were based on the full sample.

The dependent variables were the Advanced Progressive Matrices (Raven, 1962) and the Cognitive Interference Questionnaire (Sarason, 1978). The CIQ involves 11 items, each anchored on a 5-point scale with end points labelled *never* (1) and *very often* (5). The items are concerned with the extent of task-irrelevant thoughts experienced by the individual during performance on a test. Two sample items are, "I thought about how poorly I was doing", and "I thought about how often I got confused".

Appended to the CIQ were three separate line scales, each of 7 points, using the following probes and anchors: (*a*) "Please circle the number on the following scale which best represents the degree to which you felt your mind wandered during the task you have just completed" *not at all* (1) . . . *very much* (7), (*b*) "To what extent did you enjoy doing the matrices task?" *not at all* (1) . . . *very much* (7), and (*c*) "To what extent do you think this task was objectively hard or easy?" *an easy task* (1) . . . *a hard task* (7).

Procedure. Subjects were asked to complete the TAS and CFQ in tutorial classes using an imaginary name for identification purposes. In the following week each subject attended one of four scheduled workshop sessions which met in groups of between 15 and 25 students. The Advanced Progressive Matrices was administered in these sessions, held over a 2-day period. Subjects were seated at individual desks in a manner intended to replicate formal examination or group testing practice, and were asked to remain seated, and communicate only with the experimenters. At least two experimenters remained in the room throughout each testing period.

The matrices test was administered according to the standard instructions except that as part of the introductory protocol for Set I the students were informed that this was a "practice test designed to show the method of working, in preparation for the real test". Furthermore, after working for 5 minutes on Set I, and then again after 40 minutes on Set II, each subject was handed a CIQ sheet. While the subjects completed the CIQ the experimenters collected the test booklets. As in the previous session, the subjects were asked to use their imaginary names on the response sheets. Debriefing took place in the workshop sessions four weeks after the testing, because the lecture contents of their course was then concerned with the effects of anxiety, and the present findings were used for demonstration and discussion purposes. It should be noted that, in order to keep faith with the instructions given during the test, the subjects' responses on Set I were not scored.

RESULTS

In initial analyses, responses to the two assessment contexts, Set I and Set II, were compared. It was found that upon second assessment the subjects reported higher cognitive interference (means of 21·51 and 27·59), more mind-wandering (means of 2·27 and 2·81), less enjoyment (means of 4·19 and 3·29), and greater task difficulty levels (means of 3·1 and 5·11) (dependent 't' (71) values of 7·9, 3·6, 10·5 and 5·6 respectively, P<0·01 in each case).

Using the median split classifications on TAS and CFQ as independent variables, separate 2 x 2 ANOVAs were conducted on the following variables: scores on Set II of the matrices, and also cognitive interference, mind-wandering, enjoyment levels, and task difficulty levels, on each of the two assessments. It was found that the high test anxious subjects, relative to the low test anxious subjects, exhibited lower matrices scores (means of 19·65 and 24·39, F (1, 58) = 11·84, P<0·01), as well as elevated scores on the second cognitive interference assessment (means of 29·39 and 24·71, F (1, 58) = 6·11, P<0·05). Furthermore, the high cognitive failure subjects, relative to the low cognitive failure subjects, reported higher levels of interference on second assessment (means of 29·41 and 24·53, F (1, 58) = 6·7, P<0·05) and more mind-wandering (means of 3·13 and 2·17, F (1, 58) = 7·57, P<0·01). No further main or interactive effects were significant in these analyses. The breakdown of mean scores on matrices, and the second assessments of interference and mind-wandering, is shown in Table 1. An additional one-way analysis comparing upper and lower quartile groups on the CFQ upon matrices data was not significant.

TABLE 1

MEAN SCORES ON MATRICES, COGNITIVE INTERFERENCE, AND MIND-WANDERING AS A FUNCTION OF TEST ANXIETY AND COGNITIVE FAILURE

Group	N	Advanced progressive matrices	Cognitive interference	Mind-wandering
LTA-LCF	16	23·31	22·94	2·0
LTA-HCF	15	25·53	26·6	3·0
HTA-LCF	14	18·64	26·36	2·36
HTA-HCF	17	20·47	31·88	3·24
Total	62	22·02	27·05	2·66

Note: The cognitive interference and mind-wandering means refer to responses towards Set II of the matrices test. LTA-LCF refers to low test anxiety/low cognitive failure, and so on.

Correlational analyses based upon the total sample indicated several significant relationships (see Table 2). More specifically, scores on the matrices negatively correlated with cognitive interference on second assessment, and with the test anxiety questionnaire. Although CFQ scores were found to correlate significantly with cognitive interference, the correlation between CFQ and matrices was not significant. This pattern of results was further investigated using regression procedures. A stepwise regression analysis on matrices data yielded a multiple R of 0·42 which was accounted for by the test anxiety variable with no additional variables providing any further significance. A similar analysis on cognitive interference obtained a multiple R of 0·53 with significant contributions from TAS and CFQ scores but not from matrices scores. Cognitive interference also correlated with mind-wandering, and task difficulty ratings. Controlling for matrices, through partial correlation analyses, CIQ data correlated with mind-wandering ($r = 0{\cdot}66$, $P < 0{\cdot}01$, two-tailed) and task difficulty assessment ($r = 0{\cdot}25$, $P < 0{\cdot}05$, two-tailed). Mind-wandering was also correlated with CFQ scores ($r = 0{\cdot}29$, $P < 0{\cdot}05$, two-tailed).

TABLE 2

PRODUCT MOMENT CORRELATIONS BETWEEN TAQ, CFQ, COGNITIVE INTERFERENCE, AND ADVANCED PROGRESSIVE MATRICES

Group	Cognitive failure	Cognitive interference	Advanced progressive matrices
TAS	0·15	0·41**	−0·40**
CFQ		0·34**	0·05
CIQ			−0·27*

* $P < 0{\cdot}05$. ** $P < 0{\cdot}01$, two-tailed, $N = 72$.

Note: The cognitive interference (CIQ) data stem from responses towards Set II of the matrices.

DISCUSSION

After working on Set II of the Advanced Progressive Matrices, students reported significantly higher levels of cognitive interference, mind-wandering, task difficulty, and less enjoyment, than they did after working on Set I under "practice" conditions. Thus, the assumption that Set II would provide a relatively more stressful experience than Set I was supported. Further, the students' responses on questionnaire measures of test anxiety (TAS) and cognitive failure (CFQ) obtained prior to the testing session were found to be unrelated to their reported cognitive interference, mind-wandering, task difficulty, and enjoyment experienced while working on Set I. Apparently, Set I provided a relatively low-stress experience.

On the other hand, within the more stressful context of working on Set II, cognitive interference was found to relate significantly with scores from both the TAS and CFQ questionnaires. Further, the TAS was predictive of relatively impaired performance on the matrices test itself (i.e., Set II). The CFQ also predicted the degree of mind-wandering subjects reported after working on Set II.

The initial analyses, using median split groupings, were further reinforced with regression analyses which clearly established that although cognitive interference on Set II was significantly correlated with the two correlated variables of TAS and matrices scores, in a manner consistent with initial expectations and prior research (Sarason, 1978), a significant amount of the variance on the interference measure was independently attributable to the CFQ, i.e., a factor correlated with neither TAS nor matrices data. Logically, it might be suggested that variance on the interference measure could also stem from actual test performance, on the assumption that knowledge that one is doing poorly would further increase subjective distress. While in this study no assessment was made for subjects' perceptions of their performance, it can be noted that actual scores on the matrices test did not contribute towards variance on the interference measure once the more strongly predictive TAS and CFQ data had been entered into the stepwise analysis.

The initial hypothesis advanced, that TAS scores would relate to interference data on Set I, as well as Set II, was not supported. Presumably, the knowledge that Set I was "practice" was sufficient to prevent arousal of anxiety-related thinking in HTA individuals. The second hypothesis was supported and it should be noted that *both* the TAS and the CFQ predicted vulnerability to cognitive interference during formal testing. The third hypothesis, that cognitive interference would relate to matrices performance was also supported, but must be re-interpreted in light of the result that the TAS, which represents a generalised or dispositional measure, was more strongly related to matrices performance than its postulated process measure, the CIQ. Furthermore, it should be acknowledged that a certain level of interference had an apparent basis in cognitive failure, i.e., a factor *not* predictive of actual performance on the test itself.

This latter point is worthy of attention particularly since it was possible to identify a subgroup of the present sample, the low test anxiety/high cognitive failure group, who exhibited moderate levels of cognitive interference (mean of 26, see Table 1), and high matrices scores (mean of 25). In contrast, the high test anxiety/low cognitive failure subgroup exhibited a similar level of interference but considerably lower matrices scores (mean of 18). Although to do so would be going well beyond the present data, it is possible to speculate that the arousal of cognitive interference, in potentially stressful contexts, is actually helpful to high cognitive failure individuals, provided that their anxiety level remains within a manageable range. It is notable that, in the present data, high cognitive failure individuals were more likely to report the experience of mind-wandering under formal testing. The awareness of mind-wandering in the individual may, in itself, serve to stimulate additional self-cueing which, although essentially adaptive in the sense of bringing the mind back onto a task, becomes further manifest in enhanced scores on measures of cognitive interference. In other words, the occurrence of self-oriented cognitive interference may take on distinctly different functions and consequences depending upon its psychological basis, whether in the realm of anxiety or the cognitive failure characteristic.

Plainly, additional studies conducted under more controlled circumstances are needed to investigate the relationship between cognitive failure disposition, stress reactions, and possible effects upon performance in different situations. The present study confirms Broadbent *et al.*'s (1982) finding that the CFQ, as a research measure, is predictive of individual differences in psychological reactions towards naturalistic stressors, and hence adds to the construct validity of this instrument. The very human characteristic of cognitive failure, or slippage, represents a topic worthy of further study and research.

ACKNOWLEDGMENT. — This manuscript was completed while the senior author was on a sabbatical programme at the Department of Educational Psychology, University of Calgary, Alberta, and the assistance of this department in the preparation of this paper is gratefully acknowledged. Authors' address: South Australian College of Advanced Education, Lorne Avenue, Magill, S.A. 5072, Australia.

REFERENCES

BROADBENT, D E , COOPER, P. F , FITZGERALD, P., and PARKES, K. R. (1982). The Cognitive Failures Questionnaire (CFQ) and its correlates *Br. J clin Psychol.*, 21, 1-16.
CROWN, S., and CRISP, A. H. (1966). A short clinical diagnostic self-rating scale for psycho-neurotic patients. *Br. J Psychiat.*, 137, 17-25.
RAVEN, J. C. (1962). *Advanced Progressive Matrices*, 1962 Revision. London: H. K. Lewis.
SARASON, I. G. (1972). Experimental approaches to test anxiety: Attention and the use of information. In SPIELBERGER, C. D. (Ed.), *Anxiety: Current Trends in Theory and Research*, Vol. 2. New York: Academic Press.
SARASON, I G (1978) The Test Anxiety Scale: concept and research. In SPIELBERGER, C. D., and SARASON, I. G. (Eds.), *Stress and Anxiety*, Vol. 5. New York: John Wiley.
SARASON, I. G., and STOOPS, R (1978). Test anxiety and the passage of time. *J. consult. clin. Psychol*, 46, 102-109.

(*Manuscript received 26th July*, 1984)

Br. J. educ. Psychol., 55, 34–44, 1985

SCHOOLS AND SMOKING: SCHOOL FEATURES AND VARIATIONS IN CIGARETTE SMOKING BY CHILDREN AND TEACHERS

BY M. R. D. JOHNSON
(*Research Unit on Ethnic Relations,* University of Aston in Birmingham*)
B. R. BEWLEY
(*Department of Community Health, London School of Hygiene and Tropical Medicine*)
M. H. BANKS
(*MRC/SSRC Social and Applied Psychology Unit, University of Sheffield*)
J. M. BLAND
(*Department of Clinical Epidemiology, St. George's Hospital Medical School*)
AND D. V. CLYDE
(*Department of Community Medicine, St. Thomas's Hospital Medical School*)

SUMMARY. There is a need for greater understanding of influences on childhood uptake of the smoking habit. Some evidence has been put forward that 'school characteristics' as much as the personal characteristics of children and their families may play some role. A cohort study in Derbyshire secondary schools provided an opportunity to examine this thesis.

The role of school organisational features unrelated to smoking, such as are found to affect educational achievement, was generally slight and inconclusive. Other more directly related issues such as health education provided counter-intuitive results which may suggest a reversal of the causal hypothesis. Social class and peer influences were important and underline the need for a better understanding of the psychology of life-styles and role-modelling in smoking behaviour modification.

INTRODUCTION

IT has been suggested that the use of tobacco among the adult population is declining due to a greater awareness of its consequences (RCP, 1977; GHS, 1983). However, smoking by children or adolescents is still a matter of concern (Dobbs and Marsh, 1983), particularly since it is now clear that the habit or addiction begins at an early age. Indeed, this under-age group is estimated to represent a market worth some £60 million per year in Britain.

While much attention has been paid to children's attitudes and academic achievement, parental smoking and other socio-demographic variables associated with smoking by children (McKennell and Thomas, 1967; Sugarman, 1967; Bynner, 1969; Murray *et al.*, 1983a, 1983b), it is difficult to highlight those who are most 'at risk' and consequently make anti-smoking health education more effective. While it is generally agreed that the child's social environment is important, there is also a need to find out whether the school environment plays any role in affecting the likelihood of children smoking. It has been suggested that the headmaster's smoking may influence that of his pupils (RCP, 1977). Different types of schools have been found to have different prevalence rates of smoking amongst children (Holland and Elliott, 1968; Bewley *et al.*, 1973; Bewley and Bland, 1977). Power *et al.* (1972) have demonstrated significant differences between schools with regard to delinquency rates, while Banyard (1980) notes the role of the school regime in promoting delinquency. The

---

* Now at University of Warwick.

MRC/Derbyshire study provided an opportunity to examine the effect of school environment upon the child's social behaviour and to explore the apparent lack of success of anti-smoking health education (Jeffreys *et al.*, 1967; Holland and Elliott, 1968) and also to assist in its future formulation (Murray *et al.*, 1982).

Some of the variation between schools may be associated with their catchment area and therefore not susceptible to intervention, but other factors, such as staff smoking, the provision of health education, or the treatment of pupils found smoking may be easier to influence should they prove to be associated with children's behaviour. The aspects examined here may be summarised as school organisation, staff influence, health education, pressures from other pupils, and as parental influence may in some cases have an effect on the running of the school, parental background. While the ethnic composition of the school's populations was observed to vary, we were unable to collect this information, but might suggest its inclusion in a future examination of the role of home influences.

As a working set of hypotheses we sought to classify schools according to characteristics commonly used to describe them in educational research — such as pupil-teacher ratios, setting or banding procedures, organisational type (selective/non-selective, single or mixed sex, etc.), catchment area and religious foundation. No common directional hypothesis was attributed to these factors: although we were aware that urban mixed-sex non-selective schools had demonstrated higher prevalences in previous research, in this instance we felt our research could only be exploratory. However, we also sought to establish the effectiveness of health education (and smoking-related discipline) and the effects of teachers' smoking behaviour, in the belief that where the former were more rigorous smoking rates should be lower and that the latter would nullify this effect. Other factors which had emerged as significant at the individual level or in other research were tested to examine their relationship with the prior variables in case they proved to have some intervening or supervening effect. Of these, the most important was social class composition.

The study

The data on which this paper is based were drawn from a large cohort study of secondary schoolchildren in Derbyshire. The detailed methodology has already been published elsewhere (Banks *et al.*, 1978). The sampling frame was defined as all state schools in the county of Derbyshire in 1974 and schools were chosen at random in order to achieve a sample of approximately 7,000 children. In 1974 data were collected from 48 schools. Questionnaires for self-completion were distributed to all children in the 11- to 12-year age group, and an anonymous questionnaire was completed by teachers from these schools. Information about each school was completed by two members of the research team who interviewed the head teachers. Data from these three sources are used with teacher and children's data aggregated for each individual school. The prevalence of smoking was derived from the children's answers: those classified as 'smokers' had replied positively to one of the following, (1) "I have smoked sometimes but I don't smoke as much as one a week", (2) "I usually smoke between one and six cigarettes a week", and (3) "I usually smoke more than six cigarettes a week". Those described as 'regular smokers' were classified by (2) or (3). Despite the apparent fragility of this methodology, it has been copied extensively by similar studies and is now regarded as the standard technique (Dobbs and Marsh, 1983). Our own analysis of the cohort data has demonstrated sufficient consistency between years for us to have confidence in the results. Children were also asked for details of how many cigarettes they had smoked in the previous week and unreasonable answers or radically inconsistent ones excluded from analysis. Information of parental occupation was obtained from the postal questionnaire which was sent to parents of all children who completed a questionnaire.

RESULTS

School size, school type and response rates

In 1974 three of the schools were middle schools, 18 secondary modern, 7 grammar and 20 comprehensive (Table 1). The variation in size between schools was marked and ranged from 191 to 1,981 pupils. Comprehensive schools were significantly larger, but response rates and pupil-teacher ratios were similar across school types. Questionnaires were completed by 6,330 children, an 86 per cent overall response rate, while 1,333 teachers replied, a 69 per cent staff response rate.

TABLE 1

TYPE OF SCHOOL BY SIZE, PUPIL/TEACHER RATIO AND RESPONSE RATES

	Number	Average size	Single sex	Pupil/Teacher ratio	Staff response rate	Pupil response rate
Middle School	3	520	0	1 21	62%	90%
Secondary Modern	18	495	5	1 19·5	70%	83%
Grammar	7	670	3	1 18·8	64%	92%
Comprehensive	20	974	2	1 18·3	71%	86%

The variation in response rates, although not significant between school types, does present some difficulty in analysing results, in particular for staff data, as in eight schools fewer than half the staff replied. Overall it does not seem that pupil response rates were associated with those of teachers. In the four mixed-sex grammar schools the average staff response rate was 59 per cent, while children from these schools had the highest response rate — 93 per cent.

There was no statistically significant association between any of the variables analysed (except for one) and the response rate, which encouraged us to be confident in the validity of the data. An important exception was that a higher staff response rate was observed in schools where more staff reported smoking. Since it would have seemed most likely that smoking staff would be less willing to reply, and that assumption would conflict with this observation, we have assumed that our responses represent the true situation. Unfortunately, it was not possible to follow up non-responders as the teacher questionnaires were deliberately kept anonymous.

Cigarette use and the prevalence of smoking

The number of cigarettes actually smoked by children aged 11-12 was apparently small — for boys an average of 0·91 per week and for girls 0·23 a week. Only 4·3 per cent regularly smoked more than one cigarette per week (which is sometimes taken as the cut-off level for definition: Pearson and Richardson, 1978) although 16 per cent of children (21 per cent boys and 12 per cent girls) could be classed as smokers. Prevalence rates of smoking varied widely between schools — from 20 per cent to 37 per cent for boys, and 1 per cent to 27 per cent for girls.

There were also considerable variations between schools in the proportion of teachers smoking. The lowest rates were found in grammar schools for both teachers and pupils (Table 2) and highest rates in secondary modern schools. The schools which were classified as comprehensive in 1974 showed smoking prevalences close to the grand mean. Significance testing of these differences (t-test) gave no support to the hypothesis that the prevalence of smoking in grammar schools was significantly different from comprehensives. There was a significantly higher proportion of smoking in secondary modern schools by girls than in comprehensives, and the smoking behaviour of boys, girls and staff in secondary moderns was significantly different from that in grammar schools.

TABLE 2

PROPORTIONS OF CHILDREN AND STAFF SMOKING CIGARETTES (AVERAGES OF AGGREGATES) BY TYPE OF SCHOOL

	Total	Middle	Secondary Modern	Grammar	Comprehensive
Boys	21%	18%	25%	14%	20%
Girls	12%	12%	16%	9%	10%
Children	16%	15%	19%	11%	15%
Male staff	21%	21%	24%	15%	21%
Female staff	21%	27%	24%	13%	21%
Total staff	21%	24%	25%	12%	21%

The structure of schools

In examining the difference between schools, we considered matters of internal structure, such as presence of a sixth form, method of class formation (streaming, mixed ability, or banded classes) and religious/denominational bias (if any) of the school's foundation. Twenty-three schools (48 per cent) had a sixth form, only seven (15 per cent) streamed classes on a basis of 'ability' and seven had an explicitly religious foundation (Table 3).

The presence of older children in a sixth form had little impact upon smoking levels among first form pupils. Male teachers appeared to smoke less in schools with ability-streamed classes, though boys in such schools were more likely to have tried smoking — but not to be regular smokers. Female teachers in schools of explicitly religious foundation were less likely to smoke. For girls, smoking rates were remarkably constant.

It may be that the causes of variation between schools are due to factors external to the school. Data on this are hard to collect, and as a proxy head teachers were asked to describe their catchment area. Twenty-one heads categorised their catchment area as "mostly urban", seven as rural, five regarded their areas as suburban and 14 were "from a mixture of areas". Only one head defined his catchment area as a "mining community". No significant differences were found in children's smoking rates (using a series of t-tests at 5 per cent probability) between type of catchment area (as classified by head teachers), although children's smoking rates in the one mining area and in four secondary modern schools from 'mixed areas' were considerably in excess of other schools. Non-selective schools in rural areas appeared to have the lowest prevalence rates.

The influence of staff

One of the distinguishing features of individual schools is the pupil-teacher ratio. This ranged between 15·1 pupils:staff to 22·9 (mean 19·0). The lowest ratios were found in comprehensive schools, the highest in middle schools. There was, unexpectedly, some apparent relationship between the pupil-teacher ratio and smoking prevalence. The proportion of total children smoking was highest in those schools where the pupil:teacher ratio was *lowest*, and a similar relationship was found for staff smoking (using data derived from a separate but concurrent survey of staff). This finding was not expected; indeed the direction of correlation is somewhat counter-intuitive (if one accepts that higher ratios make "policing" of pupils more difficult, and may place more stress on staff). Although the (Pearson) correlation coefficients were "significant" ($r=0{\cdot}24$, $r^2=0{\cdot}06$, $P<0{\cdot}05$ for pupils; $r=0{\cdot}20$, $r^2=0{\cdot}04$, $P<0{\cdot}10$ for staff), the association is evidently weak and possibly indirect or caused by some intervening variable.

TABLE 3

SCHOOL ORGANISATIONAL ATTRIBUTES AND SMOKING BEHAVIOUR

	Total	Sixth form	No sixth form	Streamed	Un-streamed	Denominational	Non-denominational
Male staff smoking (cigarettes)	21	22	20	15	22	21	21
Female staff smoking (cigarettes)	21	23	19	22	21	15	22
Boys not smoking	43	42	45	37	44	39	44
Girls not smoking	59	57	61	60	59	59	59
Boys regular smokers	6	7	6	6	6	7	6
Girls regular smokers	3	2	3	3	2	1	3

(All figures percentages, aggregate)

The classification of children in this table is into 'regular' smokers — i.e., those smoking more than 1 per week, and non-smokers — i e , those who had *never* smoked a cigarette. This was chosen in order to illustrate the differences more clearly.

Regular smoking among children (that is, smoking one or more cigarettes per week) was not found to be so correlated. Staff smoking habits were also independently examined, and were not consistently correlated with children's smoking. There appeared to be no direct link between proportions of staff and children smoking cigarettes (see Table 4) or the headmaster's smoking habits. However, as previously suggested (Bewley *et al.*, 1979) the existence of some linkage for a proportion of the pupil population may have a wider significance for the development of smoking in later years. That paper reported in greater detail our survey of school-teacher smoking behaviour. In that we concluded that irrespective of *where* teachers smoked (in class, staffroom, or at home), if they did so then the children were aware of it. More than seeking any direct association between staff smoking and children's uptake of the habit, we are concerned at the consequences for health education of teachers' behaviour.

TABLE 4

PROPORTION OF TOTAL CHILDREN SMOKING TO STAFF CIGARETTE SMOKING

	Children		
Staff	Less than 12%	12½ to 18%	Over 18%
Less than 16%	8	4	7
16%-25%	5	4	3
Greater than 25%	4	6	7

$\chi^2 = 2{\cdot}3$, NS.

Health education

We would not have expected a great impact from any health education programmes between schools as these children could have only been exposed to them for less than a year. However, some variation might have been expected (Table 5). No significant differences were found between schools where pupils were or were not given periodic special lectures, taught about health during regular class teaching, or exposed to health education posters. Schools using a combination of all these methods had equally high smoking rates and on average smoking among both children and staff was lowest where no health education was given.

TABLE 5

HEALTH EDUCATION AND PROPORTIONS IN SCHOOLS SMOKING (PERCENTAGES)

	Lectures		In class		Posters			
	No	Yes	No	Yes	No	Yes	None	All
.le staff	21	22	16	22	22	19	19	22
nale staff	22	20	21	21	22	20	18	18
ys not smoking	43	44	47	43	45	40	48	44
ls not smoking	59	58	58	59	59	59	63	62
ys smoking > 1 per week	7	6	6	7	6	7	5	6
ls smoking > 1 per week	2	3	2	3	2	3	2	4
ys smoking > 6 per week	3	4	2	3	3	3	3	4
n smoking in front of upils at school*	36	62	22	46	39	47	12	78
mber of schools	36	12	9	39	31	17	6	6

(Numbers in tables percentages, aggregate)

* As a proportion of smokers.

The lowest proportion of male staff smoking was found in schools where no health education was given and a substantially higher proportion of male smokers smoked in front of children in schools where a combination of all three types of health education was used. This is obviously a matter for concern.

Only 13 of the 48 schools gave any specific anti-smoking education to the first-year pupils and while only one of these had more than 18 per cent of its pupils smoking, a third of those schools not giving any anti-smoking education were in the lowest third by prevalence. While it is possible to generate statistical significance from the data in Table 5, interpretation is difficult and the evidence in favour of anti-smoking education is not overwhelmingly conclusive for this age group of children.

Punishment for pupils found smoking

The deterrent effect of punishment for smoking should perhaps be considered as complementary to 'positive reinforcement' against smoking through health education. In half of the mixed-sex schools (19/38) boys and girls were treated differently. Methods of punishment varied from corporal punishment to "reasoning with the child" and in some schools the parents were informed (Table 6).

TABLE 6

PUNISHMENT OF PUPILS BY PERCENTAGE SMOKING IN SCHOOLS

	Total	(i) Corporal	(ii) Reasoning	(iii) (i) For boys (ii) For girls	(iv) Tell parents
Boys not smoking	44	44	37	43	46
Girls not smoking	59	48	58	60	61
Boys smoking > 1 per week	6	5	10	5	7
Girls smoking > 1 per week	3	5	3	2	1
Male teachers smoking in front of pupils	43	67	61	37	14
Number of schools	47	3	10	12	3

(Numbers in tables percentages, aggregate.)

It was found that smoking among pupils and staff was higher in schools where punishment was similar for both boys and girls. It may be that sterner punishment was felt to be necessary where levels of smoking were perceived as relatively excessive. Corporal punishment was used in 20 schools, although in 16 of the 38 mixed schools it was given only to boys and in only two was it used on girls. It appears that reasoning alone was effective in deterring girls, but that levels of smoking among boys were higher in schools where there was no corporal punishment. In one boys' school where children were given corporal punishment and their parents were informed, a significantly lower proportion of pupils reported that they had never smoked. Reasoning alone, however, was associated with higher levels of smoking among boys.

The home background by social class

There were differences in the social class composition of schools, based upon fathers' reported occupations. These differences were found to reflect the school's type, with a higher proportion of children attending grammar schools coming from the 'upper' social classes and more of those at secondary modern schools having parents with manual or semi-skilled occupations (Table 7). Nevertheless, even within school types, a relationship was found between social class and children's smoking and this finding is important when studying the influences upon children (see Johnson *et al.*, 1982, for a more detailed analysis of this point).

TABLE 7

SOCIAL CLASS DISTRIBUTION OF CHILDREN BY SCHOOL TYPE

	Total N	Middle (%)	Secondary Modern (%)	Grammar (%)	Comprehensive (%)
Social class I	296	5	3	15	6
Social class II	791	18	12	35	15
Social class III (non-manual)	337	5	5	9	8
Social class III (manual)	2523	48	60	31	56
Social class IV	643	18	16	9	13
Social class V	174	5	4	1	4
Total numbers	4764*	315	1123	589	2727

* N.B.: Total is less than 6,330 due to non-response and other problems of classification. Data here refer to social class grading of father's occupation. Information on type of school was missing in a few other cases.

Schools where higher proportions of children came from Social Class I homes were those with the lowest rates of children's smoking (Table 8). This was true for both boys and girls, and also for those who experimented (very low levels of smoking) or with children who were regular smokers (smoking one or more cigarettes per week). When the grammar schools, with their high proportion of parents in non-manual occupations, were excluded, the prevalence of smoking among children was still found to be correlated with the percentages of non-manual workers.

TABLE 8

PERCENTAGE CHILDREN SMOKING BY PROPORTION FROM NON-MANUAL HOUSEHOLDS

%Children from non-manual households	Boys	No. of schools	Girls	No. of schools
Less than 22%	25%	15	13%	16
22% to 32%	21%	12	13%	14
Greater than 32%	15%	14	10%	15

The influences of other pupils

There was a close correlation between other pupils' behaviour and children's smoking, and this peer group influence is important in predicting smoking uptake (Bynner, 1969; Banks *et al.*, 1981). We did not examine this in detail but observed associations between the levels of smoking among boys and girls attending the same school (Spearman's Rank test $r = 0{\cdot}51$, $P < 0{\cdot}001$). The school where boys smoked most had the third highest level of smoking among girls, and there was a large cluster of schools where neither sex smoked heavily. The correlation among mixed schools between proportions of boys and girls classed as 'regular' smokers was also significant ($r = 0{\cdot}43$, $P < 0{\cdot}005$). However, this trend was not uniform as the second and third highest proportions of girls smoking regularly (out of 38 schools) were found in schools ranking twentieth and twenty-third in proportion of boys smoking.

Power *et al.*'s study (1972) of criminal delinquency found a consistency in prevalence rankings over time, which confirmed to them the existence of a school effect. We could not undertake a completely comparable analysis, but in 1974 for a subsample of 13 schools we also collected some information from pupils at fourth form level. The evidence was not strong, but a significant association was found between proportions of boys 'never having smoked' in both first and fourth years ($r = 0{\cdot}52$, $P < 0{\cdot}05$) and between a high proportion of fourth year girls being regular smokers and a high proportion of first year girls having experimented with smoking ($r = 0{\cdot}46$, $P < 0{\cdot}06$).

DISCUSSION

This study has sought to re-examine some of the dimensions suggested by other studies as being associated with smoking. Consideration of response rates did not find any associations which would lead to serious concern about the validity of results.

Smoking levels in terms of number of cigarettes smoked per week were low, but 11- to 12-year-old children were found to be smoking enough to be classified as smokers, and to answer questions about why and with whom they "usually smoked". These proportions were very high in certain schools.

In general secondary modern schools were the most likely to have the highest proportion of smokers and grammar schools the lowest; comprehensives occupied the middle ground. This is closely associated with the composition of parental social class and at this point in the children's development data demonstrate that parental influence is an important determinant of smoking behaviour. While school re-organisation has perhaps lowered the significance of the first finding, social class composition in comprehensive schools remains far from even and this factor is regarded as highly significant. If current moves to restore selective schooling gain momentum, then this distinction may become even more crucial.

Internal differences in the organisation and structure of schools were not found to be linked with differences between the schools in smoking prevalence rates, but this may be due to the short time that the children have been there. There was some evidence here that girls' and boys' smoking differed in its relationship to potential explanatory variables which may indicate that a more sophisticated approach to health education is required. This will be complicated by the trend away from single-sex schools.

The presence of an anti-smoking education effort in the first year of school would appear to show only a slight result on the prevalence rate for first year pupils. The value and efficiency of such education — which of itself shows no impact on smoking rates — is not entirely proven. Similarly, punishment for those pupils found smoking may deter or delay pupils from experimenting or lead to a more conclusive result (whether an aversion to smoking or a defiance of school regulations). It has been shown to have some differential effect, particularly insofar as different methods of treatment may be found for boys and girls. However, we may also suggest that the effect works in the opposite direction — that education and punishment are seen as more necessary where smoking rates are highest.

Staff smoking has been described and is considered to be an important potential influence, particularly as it too varies considerably between schools. There is no causal link stipulated, and no unequivocal association with children's smoking could be detected. However, a less direct effect is suggested, through an influence on children's perceptions of adult norms and of health education incorporating anti-smoking advice.

The most important factor in child smoking habits identified by other studies has been the influence of peer-group behaviour. A school containing 'early beginners' (Rogers, 1962) may prove to develop as a school where the risk of other children smoking is highest. Limited information on fourth year children suggests a slight link between age groups, but it is expected that more influence will be found between children of the same age cohort: certainly the proportions of boys smoking relatively heavily are associated with those of girls doing likewise, even if lower levels of experimentation do not demonstrate this link between the sexes.

CONCLUSIONS

It would appear that there were considerable differences between individual schools in their smoking rates. These differences were most marked between secondary modern and grammar schools, a formal distinction that was being abolished by a policy of re-organisation into comprehensive schools. This difference would appear to be explained by differences in the social class of the pupils' parents, since within non-selective schools the prevalence rates are seen to be related to the proportions of children from higher or lower social class homes. Differences have been shown between schools in their use of anti-smoking and other health education, smoking among staff, and internal organisation, but these do not in general provide good predictors of smoking prevalence among first year pupils.

Evidently there are many characteristics of pupils (and schools) which could be used to discriminate between groups. Gender and social class background have been shown to have a relationship with reported smoking behaviour, and at this level of aggregation appear to provide the best predictor of likelihood. There are also differences between different age groups, but schools do not seem to have shown the same rank ordering among older pupils as among the first-years. It is likely that children in their first year at a secondary school will not have become completely socialised to the accepted norms of behaviour (whether from staff or senior pupils) at that school; nor will health education have had much opportunity to influence their behaviour. Consequently the lack of significance found in the use of the discriminatory indices discussed in this paper need not necessarily be considered to rule out their emergence as important factors in analysis of later behaviour. It is possible that the influence of parental social class (and inferred consequent behaviour or attitude) will decrease as the children develop. The complexity of the problem here demonstrated indicates the necessity for such research.

ACKNOWLEDGMENTS. — The study was sponsored by the Medical Research Council with the co-operation of the Derbyshire Education Authority and Area Health Authority. The research was carried out by the Department of Community Medicine, St Thomas's Hospital Medical School, London.

REFERENCES

BANKS, M. H , BEWLEY, B. R., BLAND, J M., DEAN, J. R., and POLLARD, V (1978). Long-term study of smoking by secondary school-children. *Archives of Disease in Childhood*, 53, 12-19.

BANKS, M. H., BEWLEY, B. R., and BLAND, J. M (1981). Adolescent attitudes to smoking — their influence on behaviour. *Int J. health Educ.*, 24 (1), 39-44

BANYARD, P (1980). *Delinquency — One Hundred Interviews.* Trent Polytechnic

BEWLEY, B. R., and BLAND, J. M. (1977). Academic performance and social factors related to cigarette smoking by school-children. *Br. J. prevent. Social Med.*, 31, 18-24.

BEWLEY, B. R., DAY, I., and IDE, L. (1973). *Smoking by Children in Great Britain· a Review of the Literature* London: SSRC/MRC.

BEWLEY, B. R., JOHNSON, M. R D., and BANKS, M. H. (1979). Teachers' smoking *J. Epidem commun. Health*, 33 (3), 219-222.

BYNNER, J. M. (1969) *The Young Smoker*. Government Social Survey/HMSO.

DOBBS, J., and MARSH, A. (1983). *Smoking among Secondary School children.* SS1181, SSD/OPCS, HMSO.

GENERAL HOUSEHOLD SURVEY — 1982 (1983). SS457/J. London: HMSO.

HOLLAND, W. W., and ELLIOTT, A. (1968). Cigarette smoking, respiratory symptoms and anti-smoking propaganda. *Lancet*, 1, 41-43.

JEFFERYS, M., NORMAN-TAYLOR, W., and GRIFFITHS, G. (1967). Long-term results of an anti-smoking education campaign. *Medical Officer*, 117, 93-96.

JOHNSON, M. R. D., MURRAY, M., BEWLEY, B R., CLYDE, D. V , BANKS, M H., and SWAN, A V (1982). Social class, parents, children and smoking. *Bull. Int. Union against Tuberculosis*, 57, 3-4, 258-262.

MCKENNELL, A. C., and THOMAS, R. K. (1967) *Adults' and Adolescents' Smoking Habits and Attitudes.* Government Social Survey/HMSO.

MURRAY, M., SWAN, A. V., ENOCK, V., JOHNSON, M. R. D., BANKS, M. H., and REID, D. (1982). The effectiveness of the H.E.C. 'My Body' project. *Health educ. J.*, 41 (4), 126-130.

MURRAY, M., SWAN, A. V., BEWLEY, B. R., and JOHNSON, M R. D (1983a). Some factors associated with increased risk of smoking by children *J. child Psychol. Psychiat.*, 24 (2), 223-232.

MURRAY, M., SWAN, A. V , BEWLEY, B. R., and JOHNSON, M. R. D. (1983b). The development of smoking during adolescence — the MRC/Derbyshire Smoking Study. *Int. J. Epidem.*, 12 (2), 185-192.

PEARSON, R., and RICHARDSON, K. (1978). The smoking habits of 16-year-olds in the NCDS *Public Health*, 92, 136-144.

PALMER, A. B. (1970). Some variables contributing to the onset of cigarette smoking among junior high school students. *Soc. Science Med.*, 4, 359-366.

POWER, M J., BENN, R. T , and MORRIS, J N. (1972) Neighbourhood, school and juveniles before the courts. *Br. J. Crimin.*, April, 111-132.

ROGERS, E. M. (1962). *The Diffusion of Innovations* New York. Free Press.

ROYAL COLLEGE OF PHYSICIANS (1977). *Smoking or Health.* Tunbridge Wells: Royal College of Physicians/Pitman Medical

SUGARMAN, B. (1976). Involvement in youth culture, academic achievement and conformity in school. *Br. J Sociol.*, 18, 151-164.

(*Manuscript received 5th May*, 1984)

Br. J. educ. Psychol., 55, 45-52, 1985

THE CHANGING PATTERN OF READING ERRORS AND READING STYLE FROM 5 TO 11 YEARS OF AGE

BY LEONORA M. HARDING
(The Education Centre, New University of Ulster)

JOHN R. BEECH AND WILLIAM SNEDDON
(Psychology Department, New University of Ulster)

SUMMARY. Reading errors were examined in 140 children ranging from 5 to 11 years in reading age. Intelligence, sex and social class were controlled within each reading age. Interjudge reliability on the ratings was quite high at 0·90, 0·95, 0·67 and 0·83 on graphophonic proximity, syntactic, semantic and phonemic similarity, respectively. The build-up in sight vocabulary resulted in increases in graphophonic and phonemic similarity up to 8 years of reading age. There was also a change in reading style from a phonics to a more whole-word style of reading. The more proficient readers appeared to process larger units of information which could account for systematic reductions in syntactic similarity. The reader at this level appears to be more eclectic and is capable of utilising all dimensions whenever necessary.

INTRODUCTION

THE errors a child makes reading aloud should be a rich source of data for inferring the processes involved in acquiring meaning from print. Using such data, K. Goodman (1969, 1973) implied that the less proficient reader concentrates on graphophonic cues which occur when the lettering of the child's error is similar in appearance to the original; for instance such a reader might say 'horse' for 'house'. By contrast, the proficient reader concentrates much more on semantic and syntactic cues and only uses graphophonic information as a last resort.

In a review of the research on reading errors Leu (1982) has critically examined three propositions and noted that these propositions cannot be verified by existing research. One such proposition was that proficient readers use context cues more than less proficient readers. This proposal has been confirmed by the research of Au (1977), D'Angelo (1981), Murray and Maliphant (1982) and Potter (1982), but contrary evidence is given in the research of K. Goodman and Burke (1973), Cohen (1974-75), Allington and Strange (1977), Biemiller (1977-78), Thomson (1978), Juel (1980), Stanovich (1981) and Harding (1984). The second proposal was that less proficient readers use graphophonic cues more than proficient readers. This proposal has been confirmed by the work of Clay (1969), Cohen (1974-75), Au (1977), D'Angelo (1981), but contrary evidence has been supplied by Weber (1970), K. Goodman and Burke (1973) and Allington and Strange (1977). Some of the research is contradictory because different age groups of children have been studied and alternative techniques of reading error analysis have been used. For example, Biemiller (1977-78) used 7- to 9-year-olds but did not classify errors using the Goodman reading error analysis and neither did Murray and Maliphant (1982) with a similar age group using yet another method involving a cloze procedure and an error detection task. Finally, the third proposal of Leu's was that less proficient readers should be given instruction in the use of syntactic and semantic cues to improve their reading proficiency, but this is outside the scope of the present paper.

An alternative approach to examining the patterns of reading errors of readers of differing proficiency is to examine the developing pattern of reading errors shown by children over a period of time, or to look at children at different grade levels. Biemiller's (1977-78) research provides a good example of this. In following beginning readers over one school year he outlined three phases of development. First, there was the pre non-response phase in which children made haphazard guesses at words they did not know and relied mainly on context even though they were not very good at this. Second, there was a non-response phase in which children did not make guesses presumably because they realised that there should be some graphophonic similarity between the written word and spoken word. During this phase children made graphophonic type errors. Third, there was the post non-response phase in which good readers at least became more and more proficient in the use of contextual and graphophonic cues.

Cohen (1974-75) also demonstrated a developmental trend in the types of errors shown by Grade 1 readers. Good readers it seems proceed from the non-response phase to a phase in which they substitute nonsense words (again demonstrating use of graphophonic cues) to a word substitution phase (demonstrating greater use of context). Poor readers on the other hand do not really proceed far beyond the non-response phase as they rely mainly on first and last letter cues (graphophonic).

The research by Beimiller and Cohen did not use the error analysis of Goodman (i.e., miscue analysis) as a method of measurement. The only similar study using miscue analysis was that of Y. Goodman (1971) which examined a small group of readers over only two years of reading development.

The present study aims to examine patterns of reading errors over the reading age range 5 to 11 years for the extent of their similarity along the graphophonic, semantic, syntactic and phonemic dimensions. This should provide useful information concerning changing patterns of reading. In addition, this information is compared with corresponding changes in reading style, in terms of the extent to which children use whole-word or phonic strategies in reading isolated words.

METHOD

Subjects

A total of 140 children were used in the study, including 39 already tested and analysed from Beech and Harding (1984). Basing reading age on the Schonell Graded Word Reading Test, children were selected from 5 to 11 years of reading age to give 20 within each age group. The numbers were balanced for gender and, as can be seen in Table 1, there was little variation in IQ and social class across the seven reading age groups.

Socio-economic status was determined by the father's occupation, or the mother's if the father was no longer at home. The classification was in three categories and was based on the list provided by the Registrar General (1976). Category one represented professionals, category two represented non-manual and skilled manual and the last category was the unskilled manual and unemployed.

The children were taken from five primary schools in Coleraine and Ballycastle, Northern Ireland. They were taught by the usual 'eclectic' approach, that is, when a child was first introduced to reading he or she learnt single words by the 'flash card' method. The child then proceeded through a basal reading series by reading aloud to the teacher and parents who supplied phonic and other cues as the child progressed. Phonics is also taught mainly through writing words beginning with a specified initial sound.

The other reading test used was the Neale Analysis of Reading Ability (Form 'C') which formed the basis of the reading error analysis, as it involved the children reading stories. This test contrasts with the Schonell in its basal standardised age which is 6 years compared with 5 years for the Schonell. This accounts for the apparent

TABLE 1

MEANS AND STANDARD DEVIATIONS (IN BRACKETS) OF READING AND SPELLING MEASURES, AGE, RAVEN'S SCORES AND SOCIO-ECONOMIC STATUS

	Reading age (on the Schonell)						
Measures	5	6	7	8	9	10	11
Reading — Schonell	5·5 (0·3)	6·4 (0·3)	7·3 (0·3)	8·5 (0·3)	9·4 (0·3)	10·5 (0·3)	11·4 (0·3)
Reading accuracy — Neale	6·8 (0·6)	7·3 (0·4)	7·7 (0·4)	8·6 (0·3)	10·2 (1·1)	11·1 (1·1)	11·8 (0·7)
Reading speed — Neale (words/second)	0·47 (0·16)	0·99 (0·51)	1·72 (0·66)	1·80 (0·91)	1·95 (0·94)	2·08 (1·3)	1·99 (1·4)
Comprehension — Neale	6·6 (0·3)	6·9 (0·5)	7·3 (0·4)	8·6 (0·8)	9·8 (1·3)	10·9 (1·1)	11·2 (1·0)
Spelling — Schonell	5·5 (0·3)	6·2 (0·5)	7·2 (0·7)	8·7 (0·4)	9·3 (0·5)	10·3 (0·8)	11·6 (1·0)
Chronological age	5·4 (0·2)	6·3 (0·6)	7·2 (0·9)	8·2 (0·8)	9·1 (0·9)	9·9 (0·9)	10·3 (0·7)
Non-verbal performance on the Raven's	99·3 (10·1)	101·4 (11·7)	99·2 (9·1)	101·2 (11·5)	102·4 (13·0)	103·2 (12·2)	103·0 (7·7)
Social class	1·9 (0·6)	2·1 (0·7)	2·1 (0·6)	2·0 (0·6)	2·1 (0·6)	2·2 (0·7)	2·0 (0·7)

discrepancies in reading age at the lower range in Table 1. The Neale test was standardised on Australian children, thus the Schonell was considered as a more appropriate basis for selecting reading ages as children in Northern Ireland usually start school at 4 years of age. Furthermore it is used for determining remedial pupils in the North-Eastern Education and Library Board region where this study was performed.

Two tests shown in Table 1 were administered to groups of subjects, usually a whole class at a time. These were the Raven's Coloured Matrices (revised 1978 version) and the Schonell Spelling Test 'A'. Orme's standardised tables (1970) were used on the Raven's scores and converted to non-verbal IQ equivalents. The reading speed was based on the time it took subjects to read aloud the first passage on the Neale reading test. It can be seen that reading speed doubles from the age of 5 to 6 years and doubles again by the time the child is 10.

Reading error analysis

Any errors made by the children while reading aloud the Neale were noted. These included words pronounced or accented incorrectly, substitutions of existing or new words, refusals and omissions. The first 25 errors for each subject were then transcribed, but subjects with fewer than five reading errors were omitted from this particular analysis. The transcribed errors with the original words placed next to each were then given to two independent judges who used a 10-point scale in order to rate the errors according to their graphophonic proximity, syntactic similarity, semantic similarity and phonemic similarity, respectively, as used by Goodman (1969). The level of interjudge-reliability was quite high with correlations of 0·90, 0·95, 0·67 and 0·83, respectively, across these four dimensions. The means of the two judges, for each rating dimension for each subject, were used as the basis for all subsequent reading error analyses.

Reading style test

This test is the one used in Beech and Harding (1984) and based on proposals by Baron (1979). Children had to read aloud 18 regularly spelled words (e.g., 'home'), 18 irregularly spelled exception words (e.g., 'come') and 18 nonsense words (e.g., 'bome'). Each regularly spelled word had a corresponding irregular word which had the same spelling construction in part of the word, but was pronounced differently. There was also a corresponding nonsense word for which both the pronunciations to the regular or irregular words were counted as correct. The mean word frequency of the regular and exception words were matched at 156 and 155 respectively (Kucera and Francis, 1967). The formula for the reading style index was as follows:

$$\frac{|N-R|-|R-E|}{N+R+E}$$

This formula was calculated on the median reaction times for regular (R), exception (E) and nonsense (N) words for each subject. The intention was to differentiate phonic and whole-word readers as follows. A phonics reader would find nonsense and regular words easier to read than irregular words which violate grapheme-to-phoneme correspondence rules. Therefore, (N−R) will be a smaller quantity and (R−E) will be larger resulting in a negative score in the index. By contrast, the whole-word reader's performance on the regular and irregular words will be similar, but response to nonsense words will be slower, giving a larger (N−R) and smaller (R−E) and a positive index. If subjects did not have five or more words correctly identified in any condition the index was not calculated. The words, including a 10-word practice list, were all presented on a PET microcomputer in lower case lettering by the experimenter pressing a button. A button was pressed when the child made a response and the reaction time was printed by the microcomputer.

RESULTS

Reading proficiency and reading errors

A measure of reading proficiency was taken by deducting chronological age from the Schonell reading age so that a high positive score, for instance, indicated that a child was advanced in reading for his or her age. Three separate analyses were performed for three ranges of reading age — 5 to 6 years, 7 to 8 years and 9 to 11 years of reading age. The ratings along the four ratings dimensions were correlated with the aforementioned index of reading proficiency and these are illustrated in Table 2. This table clearly shows an interesting change in reading habits as measured by reading error analysis. In the early years of reading, the good readers (that is to say those who are better than would be expected from their chronological age) produce reading errors which are *less* similar graphophonically and phonemically to the original words than the poorer readers. But a change takes place from a reading age of 7 years onwards in that the good readers in this age range produce reading errors which are *less* similar syntactically to the original words than the poorer readers.

TABLE 2

Correlations between Ratings of Reading Errors and Reading Proficiency as Measured by the Difference between Reading and Chronological Ages for Three Ranges of Reading Age

	Reading age (in years)		
Type of rating	5 to 6 (N=26)	7 to 8 (N=38)	9 to 11 (N=53)
Graphophonic	−0·34*	−0·05	0·16
Syntactic	0·32	−0·31*	−0·31*
Semantic	0·26	−0·25	−0·17
Phonemic	−0·33*	0·01	0·01

* $P<0{\cdot}05$.

Note: Children with fewer than five reading errors were omitted from the analysis.

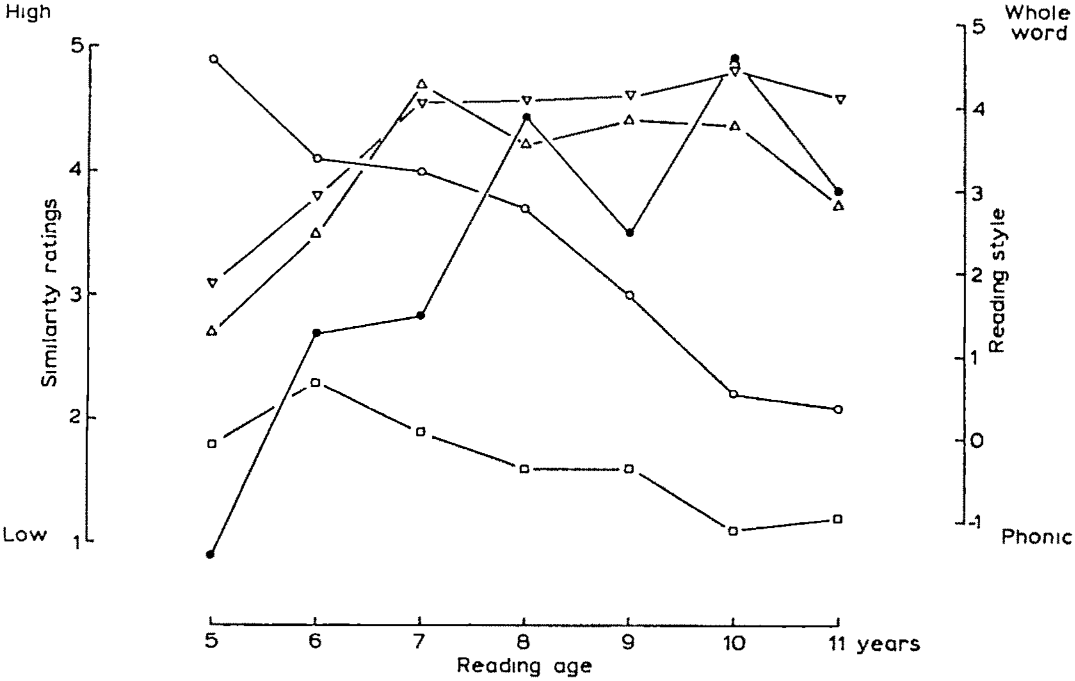
An explanation for this result could be that within the youngest age group the more proficient readers might be more adept at whole-word analysis, and the less proficient might resort more to phonics methods of word analysis. This latter group are generally the older ones within the age range and have not got off to a satisfactory start in reading. Thus when reading aloud text, the poorer readers are making errors which are graphophonically and phonemically more similar to the actual words in the text. Turning to those of a reading age of 7 years onwards, the poorer readers have entered the phase of using syntactic information, as the more proficient readers of a younger age had done, whereas the better readers have moved away from a reliance on this; perhaps in this last phase the better readers are more eclectic with their reading. They are not dependent on any particular dimension, but prefer not to use syntactic cues.

Reading errors from 5 to 11 years

The previous analysis examined reading proficiency within certain ranges. Let us now turn to changes in reading errors purely aross reading age, irrespective of whether particular readers were ahead or behind in reading. Figure 1 shows these results. The correlations between reading age and these ratings across all subjects were all significant and were 0·39, −0·51, −0·31 and 0·21 for the graphophonic, syntactic, semantic and phonemic ratings, respectively. The most systematic result was a decline in the syntactic similarity of the reading errors to the words in the text, as the children became more advanced in reading. The same result also occurred, but to a lesser extent, from a reading age of 6 years onwards for the semantic errors. By contrast, there was a weak trend towards words becoming graphophonically and phonemically more

FIGURE 1

Mean ratings of similarity indicated on the left-hand axis for graphophonic (▽), syntactic (○), semantic (□) and phonemic (△) ratings of reading errors and reading style (●) indicated on the right-hand axis as a function of reading age.



similar to the original words as reading age increased. In reality, ratings of increasing similarity occurred up to the age of 7 and thereafter reached asymptote. These two groupings were confirmed by a good positive correlation overall of 0·68 between ratings of graphophonic errors and phonemic errors and by a correlation of 0·57 between ratings of syntactic and semantic errors.

The increase in graphophonic and phonemic ratings from 5 to 7 years and then a levelling off could mean that as reading progresses a sight vocabulary is built up which means that words similar in appearance are increasingly likely to be generated; alternatively, it could mean that children increasingly use a letter-to-sound translation to decode words with which they are unfamiliar. Unfortunately a graphophonic rating can be given these two rather contrasting interpretations. The figure also indicates that there is a systematic decline in ratings of syntactic and semantic similarity over reading age. This has already been interpreted in the previous analysis as a move toward a more eclectic approach to reading.

Reading style and reading errors

Perhaps our test of reading style can throw more light on our interpretation of the results. This test produced an interesting change over reading age indicating a gradual shift from a more phonics approach to reading by a whole-word strategy as the reader matures. This change is shown in Figure 1 and was confirmed by a significant but small correlation of 0·23 ($P<0{\cdot}01$; df = 116). There was little systematic relationship between reading style and the four ratings of the reading errors. For instance, one might have expected that those using a more phonics strategy would also tend to have more phonemically similar reading errors. Instead there were no significant correlations across all subjects between the four ratings and reading styles. When examining the individual reading ages there was only one significant correlation of $-0{\cdot}47$ ($P<0{\cdot}05$) for the 7-year-olds indicating that, indeed, those 7-year-olds using a more phonics strategy also tended to make more phonemically similar errors.

When readers were grouped into the three age ranges of 5 to 6, 7 to 8 and 9 to 11 and individual reading proficiencies were taken into account, two significant correlations were found. But one of these was discounted as it was based on only 13 out of 40 readers in the 5 to 6 years reading group. The second was a low correlation of 0·24 within the 9- to 11-year-old group, revealing that those who were more whole-word in reading style tended to make less similar syntactic reading errors to the original text. This suggests a tendency for readers who are more advanced to be more whole-word in their reading and to pay less attention to syntactic cues. This concurs with the previous analyses.

DISCUSSION

Let us now synthesise these findings and integrate them with the literature, taking first those children between a reading age of 5 to 8 years. Figure 1 reminds us that over this particular period there is an increase in similarity scores along the graphophonic and phonemic dimensions of errors with the original words as well as the concomitant change in reading style from a phonics to a more whole-word approach. This result accords with some other work such as that by Weber (1970) which showed that children improved in graphic proximity scores throughout Grade 1. The more proficient readers had higher graphic similarity scores. However, as was noted before, previous research has produced conflicting results on this issue. Goodman would interpret our result descriptively in terms of our children using more graphophonic cues with reading maturity. But this is ambiguous. This could be taken to mean that the reader increasingly concentrates on the letters composing the word and translates each into its individual sound. This is probably what is usually implied by increasing graphophonic similarity. By contrast, we attribute the increasing graphophonic and phonemic similarity to the build-up in sight vocabulary, which increases the likelihood of accidentally generating words similar in appearance to the actual word. The accompanying change in reading style is another indication of this increasing sight vocabulary.

Another aspect to note about the beginning readers is a very high score in syntactic similarity which reduces linearly as a function of age. This is accompanied by much lower ratings of semantic similarity which gently decline even further. Clearly the beginning reader is making great use of syntactic information but not of semantic context. The reason for the greater reliance on syntactic information may be due to an interplay between reading speed and memory. The 6-year-old is reading slowly at one word per second (see Table 1), combined with the problems of a poor memory, this impairs the use of semantic context (Smith, 1973). As memory and reading speed improve with age, the grammatical structure of the words assumes less importance.

Turning to the older age range, this group is characterised by the same marked linear decline in similarity to the syntactic aspects of the word and a less marked decline in semantic similarity. The previous increase in ratings of phonemic and graphophonic similarity has halted by the age of 7 and thereafter remains little changed. The reading style of children in the older age range is increasingly like stage three of Gibson (1965), involving the use of "progressively higher order units of structure". They read in chunks which may be whole words or groups of words. Their reading errors tend to be more visually similar to the original than semantically or syntactically similar, hence the high graphophonic and phonemic similarity scores. This finding is not in line with previous research such as Goodman (1973, 1977) and Biemiller (1977-78). However, the decrease in syntactic and semantic similarity ratings have been obtained from a large number of subjects and on the ratings of two independent judges.

Reading error analysis has recently come in for a good deal of criticism on the grounds of non-independence of scores (Potter, 1980) and methodological weakness (Leu, 1982). We concur with much of this criticism, but believe that reading errors are a rich source of data if used carefully. We have used converging operations in making our interpretations. First, by using an index of reading style to provide supporting evidence within the same child. Second, we have monitored changes over time, using a cross-sectional design. We have inferred that across those ages in which there is a substantial change in any of our measures that these measures might well be related to one another. For example, concomitant with an increase in graphophonic similarity there was an increase in the reading style index, which enabled us to distinguish between two contrasting explanations for the increase in graphophonic similarity, involving an emphasis on phonemic awareness in young children on the one hand and whole-word reading in 11-year-olds on the other. Finally, by taking the precaution of using two independent judges, we know that the high correlations between the judges on those measures important to our analysis gives us confidence in the reliability of our results.

Requests for reprints should be sent to Dr. John Beech, Psychology Department, The University of Leicester, Leicester LE1 7RH.

REFERENCES

ALLINGTON, R. L., and STRANGE, M.S (1977). Effects of grapheme substitutions in connected text on reading behaviours. *Visible Language*, 11, 285-297.

D'ANGELO. K (1981). Correction behaviour of good and poor readers. *Reading World*, 21, 123-129

AU, K. H. (1977). Analysing oral reading errors to improve instruction *The Reading Teacher*, 31, 46-49.

BARON, J. (1979). Orthographic and word-specific mechanisms in children's reading of words *Child Dev.*, 50, 60-72.

BEECH, J. R., and HARDING, L. M. (1984). Phonemic processing and the backward reader from a developmental lag viewpoint. *Reading Res. Quart.*, 19, 357-367.

BIEMILLER, A. (1977-78). Relationships between oral reading rates for letters, words and simple text in the development of reading achievement. *Reading Res. Quart.*, 13, 223-253.

CLAY, M. M. (1969). Reading errors and self correction behaviour. *Br. J educ Psychol.*, 39, 47-56.

COHEN, A. S. (1974-75). Oral reading errors of first grade children taught by a code emphasis approach. *Reading Res. Quart.*, 10, 618-650.

GIBSON, E. J. (1965). Learning to read. *Science*, 148, 1060-1072.

GOODMAN, K. S. (1969). Analysis of oral reading miscues: applied psycholinguistics. *Reading Res. Quart.*, 5, 9-30.

GOODMAN, K. S. (1973). Miscues· windows on the reading process. In GOODMAN, K. S. (Ed.), *Miscue Analysis. Applications to Reading Instruction.* Urbana, Ill.: NCTE.

GOODMAN, K. S., and BURKE, C. I. (1973). *Theoretically Based Studies of Patterns of Miscues in Oral Reading Performance.* Final Report. Detroit· Wayne State University.

GOODMAN, K. S., and GOODMAN, Y M. (1977). Learning about psycholinguistic processes by analysing oral reading. *Harvard educ. Rev.*, 47, 317-333

GOODMAN, Y. M. (1971). *Longitudinal Study of Children's Oral Reading Behaviour.* Washington D.C.: ERIC/CRIER/Dept. of Health, Education and Welfare.

HARDING, L. M. (1984). Reading errors and reading style in children with a specific reading disability. *J. Res. in Reading*, 7, 103-112.

JUEL, C. (1980). Comparison of word identification strategies with varying context, word type and reader skill. *Reading Res. Quart.*, 17, 421-437.

KUCERA, H., and FRANCIS, W. N (1967). *Computational Analysis of Present Day American English.* Providence. Brown.

LEU, D. J. (1982). Oral reading error analysis: a critical review of research and application. *Reading Res. Quart.*, 17, 421-437.

MURRAY, L. A., and MALIPHANT, R. (1982). Developmental aspects of the use of linguistic and graphemic information during reading. *Br J educ. Psychol.*, 52, 155-169.

NEALE, M. D. (1966). *Neale Analysis of Reading Ability.* London: Macmillan.

ORME, J. E. (1970). A practical guide to estimating intelligence, attainment and intellectual deficit. *Acta Psychologica*, 32, 154-161.

POTTER. F. (1980). Miscue analysis· a cautionary note. *J. Res. in Reading*, 3, 116-128

POTTER, F. (1982). Use of linguistic context: do good and poor readers use different strategies? *Br. J. educ Psychol.*, 52, 16-23

RAVEN, J. C. (1978). *Guide to the Coloured Progressive Matrices.* London: Lewis.

REGISTRAR GENERAL (1976). *Socioeconomic Group and Social Class Allocations of Occupation and Employment Status Groups* London: H.M S.O.

SCHONELL, F (1976). *The Graded Spelling Test.* London· Oliver and Boyd.

SMITH, F. (1973). *Understanding Reading.* London: Holt, Rinehart and Winston.

STANOVICH, K. (1981) Toward an interactive-compensatory model of individual differences in the development of reading fluency. *Reading Res. Quart*, 16, 32-71.

THOMPSON, M. (1978). A psycholinguistic analysis of reading errors made by dyslexics and normal readers. *J. Res. in Reading*, 2, 7-20.

WEBER, R. M. (1970). A linguistic analysis of first grade reading errors. *Reading Res. Quart.*, 5, 427-451.

(*Manuscript received* 3rd *July*, 1984)

Br. J. educ. Psychol., 55, 53-60, 1985

ACHIEVEMENT OUTCOMES OF TWO READING PROGRAMMES: AN INSTANCE OF APTITUDE-TREATMENT INTERACTION

BY PETER FREEBODY
(Centre for Behavioural Studies in Education, University of New England, Australia)
AND WILLIAM C. TIRRE
(Institute for Personality and Ability Testing, Champaign, Illinois)

SUMMARY. The reading achievement test performance of sixth-grade pupils in contrasting reading programmes was examined. One programme (Matteson) is an explicitly sequenced, behaviourally oriented skills programme. The other (Scott Foresman) places more emphasis on rich language experience, and less on specific skill development. An aptitude-treatment interaction was found. Higher ability pupils in the Scott Foresman programme out-performed those in the Matteson programme, while lower ability pupils in the Matteson programme scored higher than those in Scott Foresman. Explanations and implications are discussed in terms of the differential degrees of benefit derived from explicit organisation of the learning environment by pupils with varying levels of academic competence.

INTRODUCTION

SOME time ago, Cronbach (1957) argued for the coalition of correlational and experimental psychology through the study of interactions between important aspects of the instructional environment, which he termed *treatments*, and student characteristics, which he grouped under the heading *aptitudes*. Cronbach and Snow's (1977) review of aptitude-treatment interaction (ATI) research showed, however, that a combination of trivially brief treatments, inadequate design and analysis techniques, and general uncertainty about how to study the phenomenon have all led to many disappointingly doubtful and inconsistent results, a conclusion also reached by Tobias (1981). There were, nonetheless, enough clear instances of ATI for Cronbach and Snow and Tobias to conclude that further work could yield significant findings for researchers if they focus on (*a*) the examination of the most plausible ATI hypotheses in long-duration, real-school studies, (*b*) the development of designs and analytic methods capable of handling such complex studies, (*c*) the development of a laboratory science for the analysis of aptitude tests and learning tasks, and the ATI constructs based on them, and (*d*), following Tobias (1982), the characterisation of treatments in terms of the frequency and intensity with which students process "instructional input".

The goal of the present study is to contribute to the first-mentioned effort, that is, the examination of a plausible ATI in a real-school setting, as a result of long-term instructional programmes. Those to be examined are two fourth-grade to sixth-grade reading programmes. The relevant aptitude is represented by two scores — an overall (verbal plus non-verbal) intelligence test score, and a reading pretest score given two years prior to the testing session, immediately prior to the pupils' entry into the separate reading programme.

The ATI hypothesis to be considered may be stated in general terms as follows. Instructional methods differ in the amount of information organising and processing effort they demand of, or provide for, the learner. The greater the information organising and processing burden *on the learner* in a given programme, the more the high-academic ability student benefits, and the less the low-academic ability student gains. Conversely, the more the instructional format *relieves the learner* of organising and processing, the smaller is the difference in the performance of high *versus* low-academic ability students (Snow, 1978). This issue has often appeared in educational debate in varying terminology — the optimal degree of structure, the benefits of discovery *versus* traditional *versus* programmed methods, and so on. All these terms, including information organising and processing burden, gloss over the issue of precisely where the critical differences lie. They may reside largely in differences in the degree to which strategies are trained, sequenced, or accessed for particular tasks, or in the degree of explicitness of expected responses, or in the amount of clear signalling of relevant knowledge. It has also been suggested (e.g., by Corno *et al.*, 1981) that a critical mediating relationship is that between structure and student anxiety levels.

Some evidence has accumulated which is compatible with the general hypothesis framed above. Researchers examining general achievement (Corno, 1979; Corno *et al.*, 1981; Sharps, cited in Cronbach and Snow, 1977), the teaching of arithmetic procedures (Thiele, 1938), instruction in introductory probability theory (Greeno and Mayer, 1976), and early reading achievement (Guthrie and Tyler, 1978) have reported data supporting the general conclusion that more systematic, highly structured, explicitly sequenced and directed instruction tended to enhance the performance of pupils rated lower in academic ability, while programmes permitting more flexibility in their learning proved more helpful to higher-rated pupils.

Specifically in the area of reading, Guthrie and Tyler (1978) showed that systematic phonic instruction led to superior early reading performance for poor readers. No balancing suppressive effects for good readers, however, were found. Bond and Dykstra (1967) reported results from the massive Co-operative Reading Program study of first-grade reading programmes. They reported no discernible ATI effects. Some questions about sample groupings and data analysis, however, led Cronbach, and Lo (cited in Cronbach and Snow, 1977) to re-analyse portions of the data. In these re-analyses, some ATI effects were found. Language-experience programmes, for instance, led to more striking differences between high- and low-academic ability girls than did whole-word drill approaches. The fact that this relation did not hold for boys illustrates Cronbach and Snow's conclusion that, while several results relating to reading instruction were suggestive of ATI, the effects were by no means convincing.

The aim of this study is to examine the possibility of ATI, at a broad level, in two reading programmes. Current programmes differ in the degree to which they explicitly sequence and structure tasks. Some explicitly train, sequence, and maintain skills and knowledge such as letter-sound correspondence, new word meanings, disambiguation skills, cross-sentence integration, and inferencing. We might hypothesise, then, that in the comparison of two long-term reading programmes, lower-academic ability pupils will show better performance in the one which provides more explicit and detailed direction than they will in less directive programmes, and that higher ability pupils will show the reverse trend.

METHOD

Subjects

The pupils tested comprised all those sixth-graders in a large school district in the greater Chicago area who had been in their present reading programme for two years or longer (N = 180). The area is, for the most part, middle-class residential. There were 96 eligible students in the Matteson programme and 84 in the Scott Foresman programme.

The reading programmes

The Matteson Four-Dimensional Reading Project. The Matteson Programme is a behaviourally oriented skills programme. As stated in the Teacher's Guide

(Schlenker, 1978) it includes: (*a*) a list of behavioural objectives for grades three through eight, (*b*) individualised learning packages covering 108 behavioural objectives in the major skill areas (word recognition, vocabulary development, literal comprehension, interpretive comprehension, analytical comprehension, study skills, and composition), (*c*) criterion-referenced pretests, post-tests, and subskill tests for each level of the programme, (*d*) individual pupil progress charts and classroom progress charts, and (*e*) commercial materials to be used with learning cards.

The programme entails branching and repetitions when test scores dictate. Some examples of the behavioural objectives are:

(1) Word discrimination: given a list of 44 Dolch sight words, the student will pronounce correctly the words.
(2) (47) Analogy: given a pair of words that are analogous, the student will identify correctly another pair of words that are related in the same way.

The Matteson programme is strictly sequenced and closely monitored. In Snow's (1977) terms, it seems to be one which, relative to others, relieves the learner of many difficult and unspecified organisational and processing tasks.

The Scott Foresman Reading System. While the Matteson programme includes in its overview some general suggestions for free reading and supplemental language experiences, the Scott Foresman system has as its *central* goal the pupil's interaction with the concepts and ideas put forth by the author. The emphasis is on rich language experience and frequent discussions aimed at the literal and inferential significance of discourse. A paragraph from the introduction to the Teacher's Guide to Level 13 (*Windows*, 1972, p. 1) gives the flavour of the place of specific skills in this programme:

> "For the few pupils who may need detailed guidance as they read, suggestions are given within particular selections. Other guidance points up a specific skill or gives help with concepts and words or phrases that may be unfamiliar. The latter might be used with pupils who read independently as well as those who read under guidance."

After each passage, pupils complete test questions which relate to vocabulary and literal and inferential comprehension, but the programme itself contains no strict sequencing or criteria for progress.

Two points need to be made here. First, as already stated, both programmes entail both skill-building and rich language experience. It is at the level of the central goals that clearly differing emphases can be detected. These differences relate to sequencing, feed-back to pupils and teacher, and degree of specificity of the skill and the expected performance. It is these aspects which, it is hypothesised, characterise the difference in processing burden on the learner. Second, the teacher's guides are not a direct measure of what actually goes on in the classrooms when these programmes are used. It is possible, though highly unlikely, that some teachers use the Scott Foresman programme in a strictly sequenced and monitored fashion, making very explicit the goals in each section. Similarly, the Matteson programme could be used merely as a remedial adjunct to an experientially oriented programme. These modifications would, however, be contrary to the spirit and the letter of the guides. In addition, interviews with the teachers in this district indicated that they used the programmes as the guides direct.

The ability and reading tests

Pupils' scores were obtained on the elementary level form R of the Educational Development Series (Scholastic Testing Service, 1976). Form R includes non-verbal (figure matching) and verbal (word knowledge and use) measures of ability.

The reading tests given were the appropriate level reading subscales of the Educational Development Series. The pretest used was that administered to the students at the end of grade four. The post-test was the appropriate form administered at the end of grade six to the same 180 pupils. The reading tests entailed literal, inferential, and main idea comprehension questions, in multiple choice format, following brief passages and poems. The passages consisted of narrative and expository discourse. In the sixth-grade test, for instance, there were five passages (two narrative, three expository) and one poem and 50 multiple choice questions in all.

RESULTS

Before presenting the major findings concerning the effects of these two programmes, the initial comparability of the two groups needs to be established, since the interaction between aptitude and programme may have existed at the point of entry to the programme. Any interactions between aptitude and programme on post-tests would, in that case, be artefactual.

Means and standard deviations for the two groups are presented in Table 1. The only contrasting point is the slightly larger standard deviation on ability scores among the Scott Foresman pupils. It is also important to establish the similarity of correlations among the measures for the two groups. Tables 2 and 3 show that, with two exceptions, these correlations do not differ significantly. The two exceptions are the correlations between pretest and number of years in the programmes (z = 3·539), and the correlations between post-test and ability (z = 2·197). We have no neat explanation of the first difference. The second, however, strongly suggests an interaction between aptitude and programmes on post-test scores.

Since there was a slight difference in ability variances for the two groups, the regression slopes for ability on pretest scores were computed. For the Matteson groups, the slope was 0·57 and, for the Scott Foresman group, it was 0·54. This difference was judged to be trivial. Any interaction between ability and programme on post-test scores could not, therefore, be attributable to an interaction existing prior to the instructional treatment.

TABLE 1

MEANS AND STANDARD DEVIATIONS

	Matteson group (N = 96)		Scott Foresman group (N = 84)	
	M	SD	M	SD
Ability	114·61	14·42	113·74	15·27
Pretest	23·27	13 15	24·92	12·99
Post-test	30·53	10 01	31·02	10 21

TABLE 2

CORRELATIONS FOR THE MATTESON GROUP

	1 Y I.P.	2 Sex	3 Ability	4 Post-test	5 Pretest
Years in Programme					
Sex	0·04				
Ability	0·10	−0 03			
Post-test	0·10	0 04	0·58		
Pretest	0·20	−0·02	0·62	0·61	

TABLE 3

CORRELATIONS FOR THE SCOTT FORESMAN GROUP

	1 Y.I.P	2 Sex	3 Ability	4 Post-test	5 Pretest
Years in Programme					
Sex	0·00				
Ability	0·33	0·18			
Post-test	0·18	0·20	0·76		
Pretest	0·63	0·13	0·64	0·50	

Tables 2 and 3 show that sex did not correlate highly with any other measures. In a regression analysis, sex was found not to be a significant predictor. It did not interact significantly with course or ability, nor were there any significantly different correlations between the measures when the coefficients were computed separately for boys and girls. Thus sex was dropped from the remaining analyses.

The aptitude-programme interaction was confirmed using multiple regression analysis. Post-test scores were predicted for each reading group separately from the equations presented in Tables 4 and 5. It can be seen that, for the Matteson group, post-test is predicted by both pretest and ability. For the Scott Foresman group, on the other hand, the only significant predictor of post-test performance is ability.

A significant aptitude-treatment interaction was found and is presented in Tables 6 and 7. From Table 6 it can be seen that ability, pretest, and the course by ability interaction contributed significantly to the prediction of post-test scores. In total, over 50 per cent of the variance on the post-test was predicted. Of particular interest here is

TABLE 4

REGRESSION WEIGHTS, PARTITIONING OF VARIANCE,† AND SIGNIFICANCE TESTS FOR POST-TEST (MATTESON, N=96)

Variable	B weight	%variance	F
Ability	0·35	34·08	55·81*
Pretest (gr. 4)	0·31	10·13	16·59*
Years in Programme	5·62	0·02	<1
Y.I.P. × Ability	−0·05	0·13	<1
Intercept	−16·39		
Residual		55·56	

R=0·666; F (4, 91)=18·14; P<0·01. * P<0·01.

† Hierarchical regression was used throughout, with order of entry as shown on the tables.

TABLE 5

REGRESSION WEIGHTS, PARTITIONING OF VARIANCE, AND SIGNIFICANCE TESTS FOR POST-TEST (SCOTT FORESMAN, N=84)

Variable	B weight	%variance	F
Ability	0·28	57·16	108·57*
Pretest (gr. 4)	0·08	0·03	<1
Years in Programme	−11·44	1·08	2·05
Y.I.P. × Ability	0·07	0·14	<1
Intercept	5·77		
Residual		41·59	

R=0·764; F (4, 79)=27·74; P<0·01. * P<0·01.

the course by ability interaction. Table 7 indicates this effect. Mid-points along both the ability and pretest ranges were selected, and the predicted post-test scores for Matteson (M) and Scott Foresman (SF) pupils are presented for purposes of illustration. These predicted scores were computed using the equations presented in Tables 4 and 5. It is the case that lower-academic ability pupils who perform as well as or *better* than their ability test scores would predict on the pretest, perform better in the Matteson programme. It is equally clear that higher-academic ability students who perform as well as or *worse* than their ability test scores would predict, perform better in the Scott Foresman programme. Interestingly, the pupil with about an average score on the ability test who performs poorly on the fourth-grade pretest is predicted to do better in the Scott Foresman programme, while the equivalent-rated pupil who performs particularly well on the pretest is predicted to do better in the Matteson programme.

TABLE 6

REGRESSION WEIGHTS, PARTITIONING OF VARIANCE, AND SIGNIFICANCE TESTS FOR POST-TEST (N = 180)

Variable	B weight	% variance	F
Ability	0·23	44·36	155·33*
Pretest (gr. 4)	0·32	3·29	11·52*
Years in Programme	−1·05	0·12	<1
Course	−22·48	0·22	<1
Course × Pretest	−0·28	0·33	1·16
Course × Ability	0·27	2·27	7·95*
Intercept	−0·47		
Residual		49·41	

R = 0·711, F (6, 173) = 29·53; P < 0·01 * P < 0·01.

TABLE 7

PREDICTED MEANS FOR PUPILS AT VARIOUS COMBINATIONS OF PRETEST AND ABILITY IN THE TWO READING PROGRAMMES ON THE POST-TEST MEASURE

Pretest grade 4	Verbal and non-verbal ability combined		
	90	110	130
12	22·00* (M)† 19·39 (SF)‡	26·44 (M) 28·80 (SF)	This combination not likely to occur
24	25·74 (M) 20·38 (SF)	30·18 (M) 29·80 (SF)	34·62 (M) 39·22 (SF)
36	This combination not likely to occur	33·92 (M) 30·80 (SF)	38·36 (M) 40·21 (SF)

Note. Figures are based on the pupil with the average stay in either programme: 2·6 years.

* Standard error of estimate (Sest) was 7·47 for the Matteson group and 6·59 for the Scott Foresman group.

† Matteson.

‡ Scott Foresman.

DISCUSSION

The hypothesis of interaction between ability and instructional condition is confirmed. Pupils of lower academic ability benefited from the Matteson programme, higher academic ability students from the Scott Foresman programme. We infer that this finding represents an affirmation of the general hypothesis that a programme which relieves the learner of much of the burden of processing, by assuming the acquisition, maintenance, and checking of important subskills, will be of greater benefit to less successful pupils than will a less strictly organised programme emphasising higher-order skills and more conceptual interaction with reading materials. More successful pupils, on the other hand, will benefit more from the latter programme. Thus, as Snow (1977) suggested, a highly structured programme tends to "flatten out" the relationship between performance and academic ability.

Some speculation about why this occurs is called for. An explanation may lie in the construct of "metacognition" that has been developed in recent theory (see Flavell, 1976; Brown, 1977; Wagoner, 1983). It has been suggested that ability test performance does not relate only to the amount of information possessed by a student but also to the student's ability to organise knowledge and skills. Sensitivity to one's errors while reading, knowing when to use certain strategies to enhance recall, locating gaps in knowledge while studying, and so on, are all executive processes rather than knowledge about reading *per se*.

If a large portion of the differences between high- and low-academic ability pupil is thought to entail deficits in executive functions, then a programme that takes on itself responsibility for many of these functions would lessen the difference between those of different abilities. The less successful would be helped, and it may be that those of higher academic ability, who have developed their own ways of organising their reading and learning, find the explicit and ubiquitous organisational aspects of the highly structured programme an interference. A programme which takes such executive skills for granted, on the other hand, and concentrates on language enrichment would enhance the learning of the high-academic ability pupils and deprive those of lower academic ability of the organisational supports they need.

Some predictable qualifications need to be made to these interpretations. First, while the sample is large, there may be something distinctive about this district or these pupils which would detract from the replicability of such a result. Second, the notion of *aptitude* needs closer examination. There may be certain skills not assessed by the pretests which render the samples different in an important way at their time of entrance into these instructional programmes. Third, to reiterate an earlier point, we need to know more precisely how strictly the programme specifications were adhered to. There may have been some mixing of techniques along an "experience-based/skill-based" continuum. These qualifications amount to suggestions for further research to clarify and detail the relationships found here.

These findings have clear implications for reading educators, as both types of programmes could be available to a pupil, and initial placement could depend upon careful assessment of entering ability. The obvious danger of such a policy is that this placement can become incarceration. This can be avoided by careful and frequent monitoring of reading performance and flexibility in the movement of pupils between programmes.

Researchers and evaluators of reading instruction will also find implications in these results. Much research on reading instruction is of the "horse-race" sort (see, for instance, Bowers, 1974, and the reviews by Samuels and Schachter, 1978 and Pflaum *et al.*, 1980). The overall means of two programmes are often compared. These results suggest the soundness of Cronbach and Snow's (1977) advice to regard separately the performance of pupils of different aptitudes.

REFERENCES

BOND, G. L., and DYKSTRA, R. (1967). The co-operative research program in first-grade reading instruction. *Reading Res. Quart*, 2, 5-142.

BOWERS, J. E (1974). *Final Report Identifying, Validating and Multi-media Packaging of Effective Reading Program* Palo Alto, Calif.. American Institute for Research.

Brown, A. L. (1977). Knowing when, where, and how to remember: a problem of metacognition. In Glaser, R. (Ed.), *Advances in Instructional Psychology*. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Corno, L. (1978). A hierarchical analysis of selected naturally occurring aptitude-treatment interactions in the third grade. *Am. educ. Res. J*, 16, 391-409.

Corno, L., Mitman, A., and Hedges, L. (1981). The influence of direct instruction on student self-appraisals: a hierarchical analysis of treatment and ATI effects. *Am. educ. Res. J*, 18, 39-61

Cronbach, L. J. (1957) The two disciplines of scientific psychology. *Am. Psychol.*, 12, 671-684.

Cronbach, L. J., and Snow, R. E. (1977). *Aptitudes and Instructional Methods: a Handbook for Research on Interactions*. New York: Irvington.

Flavell, J. H. (1976) Metacognitive aspects of problem solving. In Resnick, L. B. (Ed.), *The Nature of Intelligence* Hillsdale, N.J.: Erlbaum.

Greeno, J. G., and Mayer, R. E. (1976). Structural and quantitative interaction among aptitudes and instructional treatments. Unpublished paper, University of Michigan.

Guthrie, J. T., and Tyler, S. J. (1978). Cognition and instruction of poor readers. *J. reading Behav.*, 10, 57-78.

Pflaum, S. W., Walberg, H. J., Karegianes, M. C., and Rusher, S. P. (1980). Reading instruction a quantitative analysis *Educ. Researcher*, 9, 12-18.

Samuels, S. J., and Schachter, S. W (1978). Controversial issues in beginning reading instruction: meaning *versus* subskill emphasis. In Plaum-Connor, S. (Ed.), *Aspects of Reading Education*. Berkeley, Calif: McCutcham.

Schlenker, J. (1978) *Matteson Four-Dimensional Reading Project*. Matteson, Ill.: Reading Service Center.

Scholastic Testing Service (1976) *Educational Development Series: Reading Test*. Bensenville, Ill.: Scholastic Testing Service

Scott Foresman (1972). *Windows* (Teacher's guide) Glenview, Ill.: Scott Foresman.

Snow, R. E. (1977) Research on aptitudes for learning: a progress report. In Shulman, L. S (Ed.), *Review of Research in Education* (Vol. 4) Itascu, Ill.: Peacock.

Snow, R. E (1978). Aptitude processes. In Snow, R. E., Federico, P. A., and Montague, W E. (Eds.), *Aptitude, Learning and Instruction*, Vol. 1: *Cognitive Process Analyses of Aptitude*. Hillsdale, N.J. Erlbaum.

Thiele, C. L (1938). The contribution of generalization to the learning of the addition facts. *Teachers' College Contributions to Education* (No. 763). New York: Columbia University.

Tobias, S (1981). Adapting instruction to individual differences among students *Educ. Psychologist*, 16, 111-120.

Tobias, S. (1982). When do instructional methods make a difference? *Educ. Researcher*, 11, 4-9.

Wagoner, S. A. (1983). Comprehension monitoring what it is and what we know about it *Reading Res Quart.*, 18, 328-346.

(*Manuscript received* 21*st March*, 1984)

Br. J. educ. Psychol., 55, 61-64, 1985

TEACHER STRESS AND PSYCHONEUROTIC SYMPTOMS

By C. KYRIACOU
(University of York)
AND J. PRATT
(Sheffield City Polytechnic)

SUMMARY. One hundred and twenty-seven schoolteachers completed the Teacher-Event Stress Inventory (TESI) (a measure of self-reported teacher stress) and the Middlesex Hospital Questionnaire (MHQ) (which provides scores on six scales measuring aspects of clinically recognised psychoneurotic status: anxiety, phobic, obsessional, somatic, depressive and hysterical). The correlations between TESI scores and each of the MHQ scale scores were presented separately for male and female teachers. Eleven of the 12 correlations were positive and significant. These indicated an association between self-reported teacher stress and psychoneurotic symptoms which appeared to be greatest in the areas of anxiety, somatic and depressive symptoms.

INTRODUCTION

A NUMBER of studies have indicated a positive association between self-reported teacher stress and overall measures of mental ill-health (e.g., Pratt, 1978; Galloway, 1982; Tellenback *et al.*, 1983). However, one drawback from employing overall measures of mental ill-health is that they do not enable the identification of any particular pattern of mental symptoms to be made which would better inform a consideration of coping strategies (Beech *et al.*, 1982; Fletcher and Payne, 1982).

A particularly useful measure of mental ill-health which provides a profile of psychoneurotic symptoms is the Middlesex Hospital Questionnaire (MHQ) (Crisp *et al.*, 1978). The MHQ has been widely used in studies of occupational stress and mental ill-health (see Broadbent *et al.*, 1982).

The present study sought to explore the association between self-reported teacher stress using the TESI (described below) and mental ill-health using the MHQ. It was hoped that by employing the MHQ, the study would provide some indication of whether this association was particularly strong in specific symptom areas. In the light of studies of teacher stress dealing with stress reactions (e.g., Dunham, 1976; Kyriacou and Sutcliffe, 1978) it was expected that the association would be particularly strong for the somatic scale of the MHQ.

METHOD

The measures

The Teacher-Event Stress Inventory (TESI) comprises 43 items describing potentially stressful occurrences for teachers in the school. Teachers are asked to rate each event that occurred during the day on a scale ranging from 'not stressful' to 'extremely stressful' (scored 0 to 6 respectively). In the present study teachers were asked to record their stress experience in this way for three consecutive days. The total TESI score is the sum of the three completed inventories. Full details of the TESI, including a copy of the inventory, may be found in Pratt (1978).

The Middlesex Hospital Questionnaire (MHQ) was first described in 1966 (Crown and Crisp, 1966). It is a self-report inventory designed to provide scores (0-16) on each of six scales measuring aspects of clinically recognised psychoneurotic status: anxiety, phobic, obsessional, somatic, depressive and hysterical. The questionnaire comprises 48 items (eight per scale), each a symptom scored for occurrence (0 to 2).

The present study also requested background information (sex, age, length of teaching experience, post in the school, the classes taught and hours contact time teaching over the three days recorded on the TESI, size of school, percentage of school roll receiving free school meals).

The teachers were also asked to answer three general open-ended questions: 'to what extent, if at all, do you feel you suffer from stress?', 'why is it, in your view, that you do or do not suffer from stress?', and 'how do you cope with situations that may be potentially stressful?'.

Procedure

Six first schools, nine middle schools and ten comprehensive schools in the North of England were visited, and the staff addressed at a staff meeting with a view to soliciting volunteers for the study. In view of the fact that the TESI required recording on three school days, it was felt this method of contacting teachers was important in order to ensure the necessary level of co-operation.

Each participating teacher completed the TESI on each of three consecutive school days, and completed the background information, the MHQ, and the open-ended questions at the end of the third day.

One hundred and thirty-one teachers agreed to take part in the study. Four of these were excluded from the analysis as their questionnaires had not been fully completed. The final sample consisted of 90 female teachers (24 first school, 31 middle school, 35 comprehensive school) and 37 male teachers (14 middle school, 23 comprehensive school).

RESULTS

The distribution of the scores on the TESI for the total sample was as follows: scores 0-30, 78 teachers; 31-60, 24 teachers; 61-90, 17 teachers; 91+, 8 teachers. As can be seen, the distribution was markedly positively skewed, with a mean of 35·9 and a median score of 29. The means for the male and female teachers were 33·8 and 36·7 respectively (not significantly different using a chi square test for the above categories of TESI scores).

The mean scores for the female and male teachers on the MHQ scales are shown in Table 1. Also shown in Table 1, for comparison purposes, are the mean scores on each MHQ scale that would be expected for the general population. Since the MHQ scale

TABLE 1

MEAN SCORES FOR THE FEMALE AND MALE TEACHERS ON THE MHQ SCALES AND CORRELATIONS BETWEEN MHQ SCALE SCORES AND SCORES ON THE TESI

					Correlation with TESI	
	Female teachers (N=90)	General population females†	Male teachers (N=37)	General population males†	Female teachers (N=90)	Male teachers (N=37)
MHQ scales						
Anxiety	6·3	5·7	4·0	2·4	0·43*	0·36*
Phobic	4·2	4·6	2·4	2·3	0·21*	0·41*
Obsessional	6·1	6·9	5·4	6·1	0·25*	0·23
Somatic	4·5	4·5	3·7	3·2	0·32*	0·40*
Depressive	4·4	4·1	3·0	2·5	0·43*	0·34*
Hysterical	5·4	2·5	4·6	4·3	0·19*	0·29*

* P<0·05.

† With matched age profile.

scores differ depending on the sex and age of general population subsamples, the mean scores for the general population shown in Table 1 are based on a matched age profile with the female and male teachers considered in the present study (calculated from data presented by Crisp and Priest, 1971). As can be seen, the MHQ profiles of the female and male teachers here are very much in line with that expected for a sample of the general population with a matched age profile (the only small difference indicated was on the hysterical scale for female teachers).

The correlations (Pearson's r) between scores on the MHQ scales and on the TESI are shown in Table 1 for the female and male teachers. Eleven of the 12 correlations shown are significant for $P < 0{\cdot}05$.

Further breakdown of these results and analyses involving other variables (e.g., the length of teaching experience) did not reveal any other significant differences.

A frequency count of the different points made in response to the three open-ended questions was carried out. Teachers who described their experience of stress mentioned a whole range of symptoms, of which the most frequently mentioned were being unable to relax or 'switch off' after work, feeling very tense, being emotionally and physically drained at the end of the school day, and sleeplessness. Nine teachers specifically reported receiving medical attention for ill-health ascribed to stress, which included depression, nervous exhaustion, and nervous paralysis.

The most frequently given reasons for experiencing little or no stress were long teaching experience, having a stable or adaptable personality, having a good home life, being well prepared, and being a Christian. The most frequently given reasons for experiencing stress were having a personality tendency to worry or dwell on matters, insisting on standards, and being too conscientious.

With respect to coping with stress, the most frequently given responses were trying to stay calm, sharing problems with others, keeping things in perspective, avoiding confrontations, praying, being well prepared, and relaxing after work. Eight teachers mentioned using specific relaxation techniques, such as breathing exercises and meditation.

An exploration of the association between MHQ profiles and responses to the open-ended questions indicated that there was a tendency for teachers with profiles dominated by relatively high scores on the anxiety *and* somatic *and* depressive scales to report receiving medical attention and/or taking tablets to relieve stress. In addition, there was a tendency for teachers with profiles dominated by relatively high scores on the anxiety *and* obsessional scales to report particularly frequently the use of being prepared as a means of coping with stress. No other clear patterns emerged from this exploration.

DISCUSSION

The association between the MHQ scores and the TESI scores (indicated by 11 out of the 12 significant positive correlations) supports previous research in indicating an association between reported psychoneurotic symptoms and self-reported teacher stress. However, the present study extends this finding by indicating that the strength of this association may be greatest in the areas of anxiety, somatic and depressive symptoms. This may have important implications for understanding the nature of the association between teacher stress and mental ill-health, and for the coping strategies which are used by teachers and which may be particularly successful.

The range of coping strategies reported in the present study is very much in line with those reported by Kyriacou (1980) for a sample of comprehensive school teachers, although 'trying to stay calm' received more prominence here. Although some exploration of the association between MHQ profile and coping did not identify clear patterns, this may indicate that the diversity of the approaches to coping may only in part be linked to the particular nature of symptoms. Indeed, there may be an important interaction here between level of stress experienced, teacher personality, the symptom make-up of the stress reaction, and the coping strategies adopted. This is certainly worth exploring further.

REFERENCES

BEECH, H. R., BURNS, L. E., and SHEFFIELD, B. F. (1982). *A Behavioural Approach to the Management of Stress.* Chichester. Wiley.

BROADBENT, D E , COOPER, P. F , FITZGERALD, P., and PARKES, K. R. (1982). The Cognitive Failures Questionnaire (CFQ) and its correlates. *Br. J. clin. Psychol.*, 21, 1-16.

CRISP, A. H., JONES, M. G., and SLATER, P. (1978). The Middlesex Hospital Questionnaire. a validity study. *Br. J. med. Psychol.*, 51, 269-280

CRISP, A. H., and PRIEST, R. G. (1971). Psychoneurotic profiles in middle age. *Br. J. Psychiat.*, 119, 385-392.

CROWN, S , and CRISP, A. H. (1966). A short clinical diagnostic self-rating scale for psychoneurotic patients. *Br. J. Psychiat* , 112, 917-923.

DUNHAM, L. (1976). Stress situations and responses In NATIONAL ASSOCIATION OF SCHOOLMASTERS AND UNION OF WOMEN TEACHERS (Ed.), *Stress in Schools.* Hemel Hempstead· NAS/UWT

FLETCHER, B. C., and PAYNE, R. L. (1982). Levels of reported stressors and strains amongst schoolteachers: some UK data. *Educ Rev* , 34, 267-278

GALLOWAY, D. (1982). Stress, health and satisfaction in primary school teachers in New Zealand Paper presented at the British Psychological Society London Conference, London Institute of Education, December.

KYRIACOU, C. (1980) Coping actions and occupational stress among schoolteachers *Res. in Educ.*, No. 24, 57-61

KYRIACOU, C., and SUTCLIFFE, J (1978). Teacher stress. prevalence, sources and symptoms. *Br. J. educ. Psychol.*, 48, 159-167.

PRATT, J (1978). Perceived stress among teachers: the effects of age and background of the children taught. *Educ. Rev.*, 30, 3-14.

TELLENBACK, S., BRENNER, S., and LÖFGREN, H (1983). Teacher stress: exploratory model building. *J. occup. Psychol* , 56, 19-33.

(*Manuscript received 2nd June*, 1984)

Br. J. educ. Psychol., 55, 65-75, 1985

THE BEHAVIOURAL APPROACH TO TEACHING PACKAGE (BATPACK): AN EXPERIMENTAL EVALUATION

By KEVIN WHELDALL, FRANK MERRETT and MARK BORG
(Centre for Child Study, Department of Educational Psychology, University of Birmingham)

SUMMARY. An experimental study is described which was designed to evaluate the effectiveness of a new in-service package for training teachers in the behavioural approach to teaching (BATPACK). Teachers from two primary schools were observed (each one on three separate occasions) both before and after one of the groups received BATPACK training The teachers in the control group received two handouts and a seminar on the behavioural approach. Initially both groups of teachers were very similar in terms of their rates of approving and disapproving behaviour and their classes were also very similar in terms of amount of time spent on-task. Following BATPACK training, however, it was shown that the teachers in the experimental group were now displaying significantly changed behaviour compared with the control group. They were shown to be more positive and less negative in their responses to the children whilst their classes spent more time on task.

INTRODUCTION

THE effectiveness of the behavioural approach to teaching in the normal classroom has been demonstrated in a wide variety of experimental studies carried out not only in American but also in British educational settings (O'Leary and O'Leary, 1972, 1977; Merrett, 1981; Wheldall, 1981a). These studies have focused on altering a diversity of class behaviours including classroom conduct, academic performance and aspects of social interaction. In these studies, behavioural intervention programmes were set up to change the undesirable behaviour of a single child or a small group or class of children from a wide range of educational populations from pre-school children (e.g., Wheldall and Wheldall, 1981) to college/university students (e.g., Pickthorne and Wheldall, 1982).

The increased interest in the application of behavioural principles and techniques in the normal educational setting gave rise to the problem of how best to train teachers in the application of such methods. A thorough review of the literature on training teachers to use behavioural methods is provided in Merrett and Wheldall (1984). It soon became evident that the teacher could not be trained to use these techniques effectively merely by reading behavioural manuals and/or attending courses (Cook, 1975). Training in behavioural approaches must go beyond the mere provision of theoretical knowledge. The emphasis should be on training teachers *how* to apply the behavioural approach rather than teaching them *about* behavioural principles *per se* (Berger, 1980). In the absence of skills-based training, teaching *about* can be useful but can never be a substitute for teaching *how*, as Merrett and Wheldall (1982) have shown. Their study failed to show any significant differences in tutors' ratings of key teaching behaviours between student teachers who had and those who had not previously experienced a lecture course on the behavioural approach to teaching.

The Behavioural Approach to Teaching Package or BATPACK (Wheldall *et al.*, 1983) was developed by a project team based in the Centre for Child Study at the University of Birmingham, made up of Kevin Wheldall (project director), Frank Merrett (research fellow) and, in the early stages, Alan Russell (an LEA educational psychologist). Their aim was to develop a package which would train practising primary and middle school teachers in a number of behavioural skills.

BATPACK builds on and is influenced by a number of previous attempts in the United States and in Britain to design effective in-service training in the behavioural approach to teaching (reviewed in Merrett and Wheldall, 1984). It is also informed by the programme of experimental research into the behavioural approach to teaching carried out by Wheldall and Merrett and associated students in the Centre for Child Study (reviewed Wheldall, 1981a). Finally, it should be emphasised that the authors acknowledge a debt to the ideas underpinning Direct Instruction (Becker and Carnine, 1980) and the training materials produced by the Education of the Developmentally Young (EDY) project (McBrien and Foxen, 1981). We believe that the package, which pulls together various influences and refinements, constitutes an original approach to the problem as defined in the four following distinguishing characteristics of BATPACK.

(1) *BATPACK is skills based.* In BATPACK our primary aim is to change teacher *behaviour* by teaching new skills and techniques which can be implemented directly in the teacher's own classroom. By such skills we are referring, for example, to pinpointing and observing specific behaviours, positive rule-setting, effective praising, etc. Consequently, we believe that BATPACK can only be assessed by, and its success can only be measured by, changes in observed teacher behaviour.

(2) *BATPACK is school based.* BATPACK training is typically carried out in groups of about eight or ten which would usually include the whole complement of staff of an infant, junior or middle school. It is essential that the idea is sold to a head and his/her staff as something for the whole school which will be *their* system. This school-based training system provides support by instilling a team spirit whereby all members of the staff are working together and a common language for talking about children's behaviour problems.

(3) *BATPACK is contract based.* In the BATPACK 'Preview' we specify precisely, in advance and in practical terms, what is expected of the teachers and what the course tutor will provide. These points are then embodied in a contract which is agreed and signed by each course student and the tutor. The contract requires the tutor to start and finish all sessions on time and to provide all necessary materials. Course members contract to attend all sessions on time, to complete the work assignments between sessions, to try out the suggested techniques in class and to complete the evaluation sheets at the end of the course.

(4) *BATPACK is limitation based.* The limitations of BATPACK are deliberate and are made explicit. We are not attempting to provide training in behaviour therapy. We aim only to train teachers in a limited series of key behavioural teaching skills and techniques.

BATPACK is designed to be taught by a tutor who has attended a special training course. Prospective tutors must have a good working knowledge of behavioural psychology and most will be educational psychologists. At the tutor training course, which lasts about five hours, tutors receive a copy of the BATPACK training manual which contains all the instructions for running a course, including scripted unit notes, background reference material and master sheets for making some of the software required. A copy of the BATPACK VHS training cassette and a supply of tally counters are also needed to run courses.

The BATPACK course consists of six one-hour sessions called units, taught at weekly intervals. For each unit there are five behavioural objectives. Every unit is divided into five elements so that each corresponds to an objective. The first objective is always a review item, giving the opportunity for students to ask questions and for the tutor to review aspects of the earlier units. The last objective of each unit is concerned with the practical asssignments for the following week. These always include some reading and time for the student to observe and record his/her own behaviour in responding to the class. The three other elements are used to explore practical problems (identifying problem behaviours), to develop skills (pin-pointing, praising, observing) and to explain techniques (setting rules, defining work demands).

Every unit has an accompanying 'Takehome' in which an attempt is made to supply some of the theoretical material which will inform the practical skills learned in the unit and which provides the reading assignment for the week. In the last unit an attempt is made to review all the skills and techniques which have been covered and to present some successful classroom strategies tried out by teachers in a similar school.

From the outset, teachers who have attended BATPACK courses have been required, as part of the contract, to contribute to the evaluation of the package. The results of such participative evaluation have been used to develop later versions. The prototype package (Mark I) was taught in a middle school large enough to produce two random groups each of eight teachers: one the BATPACK experimental group and the other a control. The results were mixed. Teachers said that they had enjoyed it, felt that it had been worthwhile and that they had learned from it. Teachers' theoretical knowledge of the behavioural approach had improved significantly in the experimental group compared with the control group, but whilst observation of teachers' behaviour in the classroom showed changes in the predicted direction the differences between the experimental and control groups were not statistically significant.

As a result of constructive criticism many changes were made to the content and structure of BATPACK and the next version (Mark II) was taught in another middle school. Eight teachers took part, of whom it was possible to observe four working in their classrooms before and after the course. The data showed little change brought about by BATPACK in teachers' responses to academic behaviour. However, teachers' positive responses to children's social behaviours were shown to increase substantially in all four cases whilst negative responses to social behaviour decreased in all cases, overall by about 50 per cent. Once again, the course was welcomed and favourably received by all the teachers.

In response to evaluative comment from these teachers a further major revision of the package was made (Mark III). The quality of the VHS recordings was greatly improved since sound quality in the earlier versions had been rather poor. The contract was tightened up so that, in signing, teachers now undertook to co-operate in certain specific ways, e.g., by observing, recording and reporting (anonymously) on their own classroom behaviour. They were also given a record sheet on which they were to mark their progress towards the achievement of a course completion certificate.

This Mark III package was taught to 15 teachers and nursery nurses in two infant schools over a period of six weeks. Eight of the teachers were observed before and after the course. The data showed that total positive responses increased overall by 25 per cent with positive responses to social behaviour increasing by 87 per cent, whilst total negative responses decreased by 36 per cent with negative responses to social behaviour decreasing by 38 per cent. All these results were statistically significant. In addition, positive responses to academic outcomes increased whilst negative responses to academic outcomes decreased but neither of these changes was large enough to be significant statistically.

Only one teacher failed to increase the number of positive responses given to social behaviour and only one failed to decrease the number of negative responses given to social behaviour. All but one of the teachers had developed and were displaying a set of simple rules and all but two had re-arranged their rooms to ameliorate management problems. It is important to point out that another very encouraging aspect of all BATPACK courses run so far has been the very high rate of attendance (in the high 90s per cent).

Since the effects of BATPACK training are expected to be demonstrable in terms of changes in teacher behaviour and pupil behaviour, it is imperative to carry out further research studies in the normal classroom under normal conditions. The present study seeks to extend the previous evaluations by carrying out a more objective assessment of BATPACK's claims in the context of a carefully controlled experimental study in which *independent* observers were employed to record teacher and class behaviour.

METHOD

Research design

Using a before and after experimental/control group design, two groups of teachers and their classes from two primary schools were allocated to experimental and control conditions. Teachers from both groups were observed on three separate occasions before and after the experimental treatment. The experimental group received BATPACK training whereas the control group attended a seminar on the behavioural approach to teaching and received two handouts on behavioural methods.

Aims and hypotheses

The aim of the present study was to evaluate the effectiveness of the latest, revised version of BATPACK (Mark III) by comparing the increase/decrease in relevant teacher classrom behaviours and a measure of class behaviour. It was, therefore, hypothesised that, compared with the control group, teachers in the experimental group would increase the number of positive responses to both social and academic behaviours. It was also expected that the number of teacher positive responses contingent with reference to what has been done would increase. (Positive responses are 'contingent with reference to what has been done' when the teacher specifies *why* the pupil is being praised or rewarded.) We also hypothesised a decrease in the number of teacher negative responses to social and academic behaviours in the experimental group. Finally, we predicted an increase in pupil attending (on-task) behaviour in the experimental group compared with the control group.

Subjects and settings

Subjects consisted of 11 teachers and their classes from two primary schools situated in two suburbs in the West Midlands. The first group of six teachers, from one school, were assigned to the experimental condition: that is, BATPACK training. The remaining five teachers from the other school constituted the control group. It was not possible to assign subjects randomly to the two groups, nor was it possible to match subjects, but both groups, in fact, proved to be similar in various respects. With the exception of one of the experimental group teachers, who is in charge of withdrawal groups for remedial teaching in English, the two groups of regular teachers were equal in number and matched each other closely in terms of age-range taught, level (infants/primary) and type (normal/special) of class taught, class size and even of sex ratio for both teachers and children.

Observation schedule

The classroom observation instrument used (CTP2) is one of a series of three schedules developed by Merrett (1980) based on a schedule first proposed by Persons *et al.* (1976). CTP2 enables the observer to record systematically behaviours relevant to classroom management. The schedule is made up of two sections. In section A the observer is concerned with positive and negative teacher responses to academic and social behaviours and in section B with the on-task behaviour of the children. A particular feature of this observation schedule is that it distinguishes what we call REX responses within the positive teacher behaviour categories. REX refers to rule related approval (R), the use of approved behaviour as an example to others (E) and to occasions when teachers refer directly to what they are approving of (X). Provision also exists within the schedule for coding, for each separate 3-minute observation segment, information about the arrangement of the classroom and the position of the teacher but these data were not employed in the present study. A detailed description of the classroom behaviours observed and the instructions for the use of the schedule are included in the Appendix.

Procedure

The pre-course observations of the experimental group took place in the last week of January. The first of the six (weekly) one-hour sessions making up the BATPACK course was held on 31st January. The other sessions took place before the middle of March, with one interruption due to the spring half-term vacation. All meetings started at 3.30 p.m. (i.e., immediately after school) on the school premises. Post-course observations took place before the end of March.

The teachers in the control group were observed before and after the intervention, in February and April respectively. As soon as the first observations had been completed the teachers were given a paper outlining the behavioural approach to teaching. A seminar was held at the school in March during which a detailed discussion of this paper took place. Practical suggestions were made at that time for the application of behavioural techniques and the teachers were given a second paper which set out clear guidelines for the application of behavioural methods in the classroom.

Owing to organisational problems, school vacations, and the fact that all the observations were carried out by the same person, it was not possible to carry out the pre- and post-intervention observations of both groups exactly concurrently. There was a four-week gap between the experimental and the control group observations, both before and after the experimental treatment.

During the observation session, the observer located himself to one side of the classroom and looked at teacher and pupil behaviours in turn, as indicated by the schedule. Every effort was made to be as unobtrusive as possible and no contacts were made with the class teacher nor with any of the pupils. The observer in no way participated in the proceedings. If, in the course of an observation, classroom activity was interrupted, the observation was suspended until teaching resumed.

In order to obtain a representative sample of classroom behaviours teachers were observed during both morning and afternoon sessions, for different lessons and on different days of the school week. The type of classroom activities ranged from play-activities (in infant classrooms), to more formal lessons in subjects like Mathematics and English (in junior classrooms). Similar sorts of lessons were observed for each teacher for both pre- and post-training observations focusing on more academic lessons where possible, i.e., where children were engaged in specified set work. Teachers were observed only once on any single day.

During the post-intervention observation sessions with the experimental group it was possible to obtain some inter-observer agreement data. The regular observer (Borg) was joined by a colleague and they carried out simultaneous observations on four classes, taking care to synchronise the sections accurately. Both observers were very experienced in the use of CTP2 and both were independent of the course. Neither took part in the BATPACK course nor did they have knowledge of the course or its nature before the observations took place. The inter-observer agreement measure was calculated as follows: Inter-observer agreement =(number of agreements/number of agreements+number of disagreements) × 100.

RESULTS

The mean level of agreement between observers for section A was 94 per cent (range 91-97 per cent) and for section B 91 per cent (range 82-100 per cent) which can be regarded as satisfactory evidence for the objectivity of the data.

The results for each of the three observation sessions for each teacher were summed for pre- and post-training observations separately to yield two total scores for each teacher behaviour and two percentages for the on-task behaviour of each class. Given that half of each session was spent observing the teacher (i.e., 15 minutes) the total scores refer to the total number of behaviours in each category recorded in the 45 minutes of observation time. These totals and percentages were calculated for pre- and post-training observations and are shown in Table 1.

From Table 1 it can be seen clearly that pre-intervention means for all teacher and class behaviours showed no significant differences between the experimental and control groups of teachers. In other words, they were responding to their classes at about the same rates, in terms of positive and negative statements, gestures and actions and their classes were responding at similar levels in terms of attentional behaviour.

TABLE 1

DIFFERENCES IN MEAN VALUES ON THE BEHAVIOURAL MEASURES

		Experimental group N=6		Control group N=5		t-test df=9		Analysis of covariance df=1, 8	
Behaviours		Mean	SD	Mean	SD	t	P	F	P
Positive Academic	Pre	29 83	14·37	26 00	15 37	0·43	NS		
	Post	70 50	21 82	25 80	13 55	3·97	<0·01	40·00	<0 001
Positive Social	Pre	00 16	00·40	00·40	00·54	0 81	NS		
	Post	09 83	04·53	00·00	00·00	4·80	<0 001	29·76	<0 001
Total Positive	Pre	30·00	14·58	26 40	14·97	0·40	NS		
	Post	80·33	24·30	25·80	13·55	4·45	<0·01	49·83	<0·001
Positive Academic (REX)	Pre	02·66	02·80	03·40	02·88	0·43	NS		
	Post	28 33	12 86	03 40	02·88	4·21	<0·01	24.77	<0·001
Positive Social (REX)	Pre	00·16	00·40	00·20	00 44	0·13	NS		
	Post	03·66	02·65	00 00	00 00	3 06	<0·05	17·33	<0·01
Total Positive (REX)	Pre	02·83	03·12	03·60	02·70	0·43	NS		
	Post	32·00	14·31	03·40	02 88	4 36	<0 01	31·66	<0·001
Negative Academic	Pre	01·83	02·78	01·80	01·64	0·02	NS		
	Post	01·33	01·50	03·20	01·09	2·30	<0 05	5 92	<0·05
Negative Social	Pre	00·66	00·81	01·00	00·70	0·72	NS		
	Post	01·50	01·04	05·80	04 08	2·51	<0·05	4·95	NS
Total Negative	Pre	02·50	02·58	02·80	01·64	0·22	NS		
	Post	02 83	01·94	09·00	03·60	3·63	<0·01	15·09	<0·01
%On-task	Pre	75·00	07·45	74·80	03 42	0·05	NS		
	Post	84·00	02·68	68·80	05·58	5·94	<0·001	31·68	<0·001

The post-intervention measures, however, show a marked change. The teachers of the experimental group were now responding positively to both academic and social behaviour at a much higher rate than those of the control group. At the same time their disapproval was at a much lower rate to both categories of behaviour. In addition, the teachers in the experimental group were giving many more approving responses accompanied by reference to what has been done (i.e., REX responses), a practice encouraged by the BATPACK course. At the same time an improvement of 9 per cent was seen in the level of on-task behaviour in their classes whereas it fell 6 per cent in control classes. All these differences were shown to be statistically significant (two-tailed t-tests). Analyses of covariance subsequently performed on the post-test data, in which pre-test scores were used as a co-variate, substantially confirmed these results, with the exception of disapproval of social behaviour. Comparisons of these measures fell just short of statistical significance.

Turning now to the results for individuals it was shown that all teachers in the experimental group increased their use of positives in all categories. In fact, all these teachers *doubled* their previous rates with the exception of two teachers whose responses to academic behaviour (only) did not quite meet this criterion. Moreover, five out of the six teachers showed *increasing* use of positives over the three post-test observation sessions. In the control group only a few minor increases in positive behaviour were in evidence: they were limited to marginal increases to academic behaviour only, and were seen to be *decreasing* over the post-test observation sessions. On-task behaviour was seen to increase in five out of six experimental classes, with the sixth class maintaining its already high on-task level (81 per cent). In the control group on-task behaviour fell in all classes except one, which increased 5 per cent.

All the teachers who took the course completed the BATPACK questionnaire anonymously and a summary of these data suggests that all of them regarded BATPACK more favourably at the end than at the beginning, except two who were favourably disposed at the outset and did not change their stance. All were prepared to recommend the course to colleagues. All believed that their own responses to children had improved and that they had increased their rates of positive responding. They were less sure about the levels of their negative responses. Some of the teachers reported that they had attempted to manipulate antecedents; for example, four of them had set up rules in their classrooms. Generally speaking, the comments about the structure of the course were favourable, suggesting that the Mark III version, which has been the result of modification in the light of previous teacher comment, is beginning to approximate to a satisfactory form. Finally, attendance at the course, as is usual with BATPACK, was very good; one teacher only was absent for one session. All this suggests that teachers find the package acceptable and feel that they gain much from it.

DISCUSSION

The data obtained from the schedules showed that prior to the BATPACK training the two groups of teachers did not differ significantly, and, in fact, were very similar in terms of their use of approval and disapproval. Rates of disapproval to both academic and social behaviour were very low in both groups as was approval of social behaviour. Approval for academic behaviour, however, was (relatively) very high. Average on-task behaviour of the classes was around 75 per cent in both groups, again quite high.

Both groups of teachers in the present study differed markedly from those observed previously. Wheldall and Merrett (1985) have presented preliminary results on 'natural rates' of responding of British teachers based on 25 primary classes observed using a similar schedule. Compared with the average for these classes, teachers in both groups in the present experiment praised and disapproved appreciably less. For example, the mean rate for positives (for academic and social behaviour combined) was less than half that found previously whereas the equivalent rate for negatives was less than one-tenth that of teachers in the 'natural rates' study. The mean on-task behaviour for both groups in this experiment, however, was very high: 10 per cent higher than in the 'natural rates' study. We are confident that the observed rates are accurate reflections of the initial behaviour of the teachers in this experiment. Two highly trained and experienced observers were employed. Both had recorded more typical rates of responding (as found in the earlier study) in their observations of classes outside this experiment and inter-observer agreements between them were high. We can only conclude that both schools in this experiment were not typical. It is possible that with such high rates of on-task behaviour there is little room for negative comment from the teachers. There was, in a sense, no 'need' for interventions to reduce disruptive behaviour.

Nevertheless, and in spite of this, the on-task behaviour in the experimental classes was shown to increase as a result of the course whilst that of classes in the control group declined. The classroom behaviour of the control group teachers remained at these atypical low levels despite the fact that they had read about the positive methods advocated by the behavioural approach, discussed these at length with an enthusiastic advocate and had received a paper setting out simple ways of applying them. We are not suggesting that the control treatment was comparable with BATPACK but it is more typical of the behavioural input which teachers commonly receive, if any. It is yet another example of the ineffectiveness of learning *about* techniques rather than having training in *how* to apply them.

Examination of Table 1 shows that considerable changes were recorded in the classroom behaviour of the teachers who took part in the BATPACK course. They gave more contingently positive responses to their children and they used fewer negatives than before. Moreover, they related their contingently positive responses more frequently to the behaviours they wanted to encourage, i.e., they used what we have called REX responses more frequently. This practice clearly has didactic value and is explicitly recommended in BATPACK, especially in relation to a well-defined rule structure. Of the positive responses given by the teachers in the experimental group before the BATPACK course, 10 per cent were REX responses. After the course 41 per cent of the positive responses to academic and 36 per cent of those to social behaviours were REX responses. At the same time changes were observed in the behaviour of the classes taught by the experimental group of teachers: the children spent more time working. Although the teachers reduced the number of negative responses they gave to social behaviour the difference in response rates (pre and post) was not great since for all the teachers in this study negative responding was at a very low rate from the outset. It is important to stress that praise rates were not only higher than previously but also appreciably higher than the average levels recorded in the 'natural rates' study. The gains in praise rates in percentage terms were much greater in this experiment than in the previous evaluation studies of BATPACK Mark III. This appears to be a direct result of the initial atypical low rates of responding, i.e., these teachers had more room for improvement.

It is of interest to note that in the classes of the teachers of the experimental group, where rates of positive responding rose and those of negative responding fell, the mean on-task rates for the children rose. For the control group classes, where rates of positive responding fell (slightly) and those of negative responding rose, the mean on-task rates fell. This gave a mean difference of 15 per cent in on-task levels between the two groups of classes at the end of the evaluation experiment. The numbers involved were, of course, small but the results are in accord with the expectations of applied behaviour analysis.

This experiment has demonstrated that BATPACK (Mark III) training achieves its objectives of changing teachers' behaviour in the predicted directions and that this leads to improved classroom behaviour. More generally, this study emphasises that skills-based, in-service training is an effective means of passing on the benefits of the behavioural approach to practising classroom teachers. The potential of the behavioural approach can now be realised by training teachers both more effectively and more efficiently.

REFERENCES

Becker, W. C., and Carnine, T E (1980). Direct Instruction. an effective approach to educational intervention with the disadvantaged and low performers. In Lahey, B , and Kazdin, A. (Eds), *Advances in Clinical Psychology*, Vol 3. New York· Plenum.

Berger, M. (1980). Behaviour modification and teacher education. *Psychology Teaching*, 8, 43-48

Cook, J. (1975). Easing behaviour problems *Special Education — Forward Trends*, 2, 15-17.

McBrien, J., and Foxen, T (1981). *Training Staff in Behavioural Methods· The EDY in-service course for mental handicap practitioners (Instructor's Handbook)*. Manchester: Manchester University Press.

Merrett, F. E (1980). The application of behaviour modification in the primary school classroom. *Collected Original Resources in Education*, 4, 11 fiches.

Merrett, F. E. (1981). Studies in behaviour modification in British educational settings. *Educ Psychol.*, 1, 13-38

Merrett, F. E., and Wheldall, K (1982). Does teaching teachers about behaviour modification techniques improve their teaching performance in the classroom? *J. Education for Teaching*, 8, 67-75.

Merrett, F. E , and Wheldall, K. (1984). Training teachers to use the behavioural approach to classroom management — a review. *Educ. Psychol.*, 4, 213-231.

O'Leary, K. G., and O'Leary, S G. O (1972, 1977). *Classroom Management: The Successful Use of Behaviour Modification (First and Second Editions)*. New York. Pergamon Press.

Persons, W. S., Brassel, W. R., and Rollins, H. A. (1976) A practical observation procedure for monitoring four behaviours relevant to classroom management *Psychology in the Schools*, 13, 64-71.

Pickthorne, B., and Wheldall, K. (1982). A behavioural approach to teaching subsidiary physics to engineering students. *Educ. Psychol* 2, 79-85.

Wheldall, K. (Ed.) (1981a). *The Behaviourist in the Classroom Aspects of Applied Behavioural Analysis in British Educational Contexts*. Birmingham· Educational Review Publications

Wheldall, K (1981b). Applied behavioural analysis in British educational contexts (with particular reference to the Birmingham group). In Horne, D. J. de L., and McMurray, N. E. (Eds.), *Behaviour Therapy in Australia Proceedings of the Third Australian Conference on Behaviour Modification, 1980* Melbourne: Australian Behaviour Modification Association. Also in *British Psychological Society Education Section Review*, 5, 17-25 (1982).

Wheldall, K., and Merrett, F E. (1985) Training teachers to be more positive: the behavioural approach. In Bennett, S. N., and Desforges, C. W. (Eds.), *Recent Advances in Classroom Research*. Edinburgh: Scottish Academic Press in association with the British Psychological Society.

Wheldall, K., Merrett, F. E., and Russell, A (1983) The Behavioural Approach to Teaching Package. Birmingham. Centre for Child Study, University of Birmingham

Wheldall, K , and Wheldall, D. (1981) School meals, praise and contingent dessert. an attempt to improve eating behaviour in the nursery class. *Behavioural Approaches with Children*, 5 (2), 30-46.

(*Manuscript received 2nd June*, 1984)

CLASS TEACHER/PUPIL OBSERVATON SCHEDULE (CTP2)

PROCEDURE FOR USING THE OBSERVATION RECORD SHEET CTP2

A complete observation session will take 30 minutes. CTP2 is meant to be used when the group is engaged in classroom activity either as a whole, in groups or individually. The observer should position himself at one side of the classroom and should not interact with the pupils or the teacher in any way. The observer will need a stop watch or a watch with a sweep hand. As far as possible pay attention only to the immediate stage of recording. Try to be as objective as possible, remembering that reliability is the chief aim — that any person observing what you observe would record what you have recorded. It may be necessary to move at some stage so that the faces of all the children can be seen. This movement, like the rest of the observing and recording process, should be as unobtrusive as possible. The recording need not be continuous. If it is interrupted, start again, as soon as convenient, at the beginning of a 3-minute section. Otherwise, 3 minutes is to be spent on each section alternately. If the date is even, start with section B and if it is odd start with section A.

Section A is concerned with teacher behaviours (observation time 3 minutes). During this time the observer pays attention only to the teacher. Teacher behaviours recorded as positive events will include verbal praise (Correct! That's great! I like that! Well done! and so on); gestures like nodding encouragingly, smiling, giving the thumbs up sign; actions like placing the hand on the shoulder, the granting of privileges and the giving of tokens of approval like stars, ticks, points and so on.

Teacher behaviours to be recorded as negative events would include verbal criticism (like a reprimand), the pointing out of failure, error or general disapproval (That's wrong! You're a dead loss! Don't do that!); gestural responses like frowning or glaring; aversive contact involving shaking or smacking; withdrawal of privileges, points or rewards; isolation from the rest of the group and so on.

Other statements made by the teacher, which will be mainly concerned with informing or instructing, should not be recorded. The observer should mark the next number in the appropriate box whenever a relevant teacher behaviour is observed. For each 3-minute segment of observation in section A, place a symbol in the marginal box for teacher interaction and one for the class formation relating to the situation at the moment that segment begins.

Symbols to be used in the marginal box

TC — teacher interacting with the whole class
TG — teacher interacting with a group
TI — teacher interacting with an individual
TID — teacher interacting with an individual at teacher's desk
1 — children sitting at desks/tables arranged in groups
2 — children sitting at desks/tables arranged in rows
3 — children grouped around teacher (sitting or standing)
4 — children grouped around demonstration bench or table

Symbols to be used in segment boxes

O = event appropriately contingent
Ø = event non-contingent
R = event has been related to a rule, e.g., "Great! You have kept the rule and put your hand up without calling out."
E = event has been used as an example, e.g., a child's work or action is commented upon or displayed as an example to be copied.
X = event appropriately contingent with reference to what has been done, e.g., "Well done! That was very neat work."

Section B is concerned with child behaviours (observation time 3 minutes)

Divide the class, by eye, into three convenient and approximately equal groups and pay attention to each in turn for 1 minute. During the first minute look at each pupil in the first group, in turn, for 4 seconds and decide whether, *for the whole of the time*, he/she is on-task. To be rated on-task the child must be attending, i.e., in eye contact with the teacher or the task. This would include listening to the teacher or to some apparatus operated by the teacher or to another child asked by the teacher to speak or following instructions or routines set by the teacher, e.g., reading, drawing and so on. If the first child observed is on-task place a dot in the upper/left half of the first block. If he is off-task place a dot in the lower/right half. Continue so for each child in your first group and then wait till the first minute is up. Later it will be possible to enter the number on-task and the total number of children observed in each minute. As the second minute begins pass to the second group of children and observe them in order. The sixteenth block is for the grand total of results.

Try to ensure that all headings are completed immediately after the observation session ends.

Observation Record Sheet (CTP2)

School .... Class ....

Teacher .... Date.................... Time ....

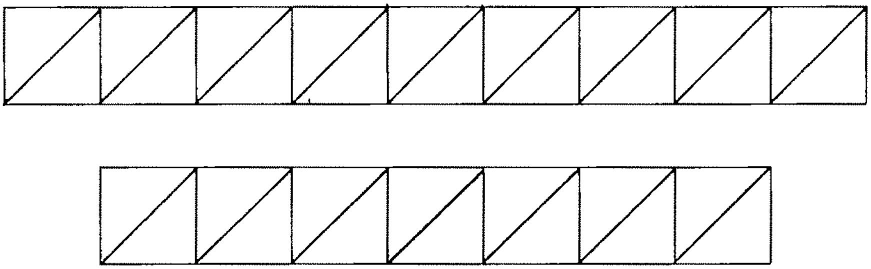
No. of children present......................... Observer ....

Nature of lesson .... Second observer.................................... ....

Section A.

	Positive events		Negative events	
	Academic	Conduct	Academic	Conduct
	1 2 3 4 5 6 7 8 9 10	1 2 3 4 5 6 7 8 9 10	1 2 3 4 5 6 7 8 9 10	1 2 3 4 5 6 7 8 9 10
	1 2 3 4 5 6 7 8 9 10	1 2 3 4 5 6 7 8 9 10	1 2 3 4 5 6 7 8 9 10	1 2 3 4 5 6 7 8 9 10
	1 2 3 4 5 6 7 8 9 10	1 2 3 4 5 6 7 8 9 10	1 2 3 4 5 6 7 8 9 10	1 2 3 4 5 6 7 8 9 10
	1 2 3 4 5 6 7 8 9 10	1 2 3 4 5 6 7 8 9 10	1 2 3 4 5 6 7 8 9 10	1 2 3 4 5 6 7 8 9 10
	1 2 3 4 5 6 7 8 9 10	1 2 3 4 5 6 7 8 9 10	1 2 3 4 5 6 7 8 9 10	1 2 3 4 5 6 7 8 9 10

Section B.



RESEARCH NOTES

Br. J. educ. Psychol., 55, 76-80, 1985

THE VALIDITY OF A SHORT FORM OF THE COOPERSMITH SELF ESTEEM INVENTORY

BY J. GIBBS
(*School Psychological Service, Barnet LEA*)
AND B. NORWICH
(*University of London, Institute of Education*)

SUMMARY The short form of the Coopersmith Self Esteem Inventory designed by Lee (Argyle and Lee, 1972) was subjected to several tests of validity and internal consistency. The short form was administered to 41 children in off-site units for persistent school non-attenders and 41 matched controls (average age 14 years 9 months). Factor analysis revealed no general factor of self esteem. Scores were found not to correlate significantly with general verbal abilities nor word reading. However, there was a moderate correlation between the Coopersmith scores and repetory grid measures of self esteem. Children who were rated by their teachers to have "neurotic" compared with conduct or no difficulties showed significantly lower scores on the Coopersmith inventory, suggesting that the short-form inventory was assessing some aspects of self perceptions and self evaluations about emotional reactions. The results are interpreted as illustrating the need for improved test construction and validation in this area.

INTRODUCTION

Despite considerable interest in self referent constructs — self concept and self esteem — there have been persistent methodological problems in assessing these constructs (Wylie, 1961; 1974; Wells and Marwell, 1976). These problems are associated with a lack of theoretical development and inadequate test construction and validation programmes (Wylie, 1974). Assessment techniques and tests embody theoretical assumptions about self processes which are often lacking in clarity. The distinction between descriptive and evaluative aspects of self processes is widely recognised but does not often influence test construction. Coopersmith (1975), amongst others, for example, distinguishes between self concepts, images a person forms of him/herself, and self esteem, positive and negative feelings and attitudes about the self. However, the names and contents of certain tests do not often reflect such theoretical distinctions, e.g., Piers Harris (1964) "self concept scale": "How I feel about myself". In relation to these issues, Anderson (1981) has underlined the importance of specific definitions of self referent constructs and has advocated a domain referenced approach to operational definitions as a means of clarifying and validating the interpretations of assessment.

Self referent constructs have traditionally referred to global enduring characteristics (Wylie, 1961) rather than to particular areas of experience. With the exception of Shavelson's hierarchic model of self concept (Shavelson *et al.*, 1980) the theoretical assumption of a unitary construct has been accepted without much prior empirical testing. Though certain measures of general self concept can be scored for specific self concepts (e.g., the academic areas of the Piers Harris self concept scale and Coopersmith Self Esteem Inventory), these procedures are not without difficulties. These subscales have been found not to differentiate between criterion groups in the same way as specific measures of academic self concepts (Chapman and Boersma, 1980) This would be expected considering that these measures were not designed to assess particular areas of self concepts comprehensively.

This research note is concerned with several tests of the validity and internal consistency of a short form of the Coopersmith self esteem inventory (SEI) designed by Lee (Argyle and Lee, 1972). The full scale SEI for 10-16 year-olds consists of 58 items to be rated as like me — unlike me. Coopersmith (1967) selected the items from a scale constructed by Rogers and Dymond, but no reference to this scale was reported. Five psychologists sorted the items into high and low self esteem groups but no indication was given of whether these psychologists shared Coopersmith's formulation of self esteem as a general enduring "personal judgement of worthiness expressed in the attitudes the individual holds towards himself" (Coopersmith, 1967, p. 10). Though Coopersmith investigated aspects of the reliability and validity of the SEI, for example, retest reliabilities of 0·88 for five-week and 0·70 for three-year periods, there were no internal factor or item analyses, nor did he investigate relationships across a wide range of SEI scores. Wylie (1974) concluded that the psychometric data were inadequate and recommended that the SEI was not an instrument of choice without further study.

Bagley *et al.* (1979) commented on the US psychometric studies favourably without referring to Wylie's evaluation. In their English study two main factors were found — termed "I'm a failure" and lack of social confidence. From this analysis 23 items with the highest loadings were selected for a short form. A principal components analysis of this short form obtained different loadings, a finding which Bagley attributed to different statistical techniques. The internal structure for the short form was judged to have satisfactory validity for both sexes on the grounds that 19 items had loadings above 0·20. Tests of validity revealed moderate correlations between the full SEI and a measure of anxiety, and lower correlations (0·2 to 0·3) with IQ, reading level, sociometric status and Cattell's measure of ego strength. The short form of the SEI (25 items) reported in this study was based on an unreported item analysis of the full scale conducted by Lee in the UK. It is interesting that only 10 items are common to both the Lee and the Bagley short forms (23 items). The Lee short form and the full SEI correlated 0·86. Reducing the scale from 58 to 25 items also lowers the internal consistency of the scale.

It was decided therefore to examine the internal consistency of the Lee short form and its factor structure in order to determine whether it assesses a global factor. Its construct validity was examined by testing (1) its relationship with verbal abilities and word reading skills, both of which would be expected to be weakly associated with self esteem (Coopersmith, 1967), and (2) its relationship with emotional and behavioural problems. It would be expected that children with predominantly neurotic anxiety problems would have lower self esteem than children with conduct or no problems (Coopersmith, 1967; Rosenberg, 1965). Concurrent validity was examined by testing the relationship between the short form SEI measure and an alternative self report measure of self esteem

METHOD

Sample

Forty-one children were selected from three off-site units for persistent secondary school non-attenders. The average age was 14 years 9 months, ranging from 13 years 5 months to 15 years 10 months. Another 41 children individually matched to the first group for age, sex and social class, were selected from local comprehensive schools feeding the units. There were 48 boys and 34 girls in the total sample.

Measures

Test data were collected as part of a wider study of non-attenders, with each child individually tested by an educational psychologist on the following measures:

(1) *Verbal intelligence.* Most children performed the five verbal subtests of the Wechsler Intelligence Scale for Children — Revised. Fourteen children from one unit were tested on the five subtests of the unrevised WISC.

(2) *Reading· word recognition.* The Holborn reading scale (Watts, 1948), though dated, was used as a measure of word recognition in a sentence context The scale which provides reading ages from 5 years 9 months to 13 years 9 months was selected because of its ease of administration and because it assesses word recognition in context.

(3) *Self esteem.* The Lee short form used in this study was scored so that high scores represented low self esteem and low scores represented high self esteem. This was in the opposite direction to that used by Lee but this makes no difference to the analyses of the short form.

In addition to the Lee short form of the Coopersmith self esteem inventory described above a second measure of self esteem was derived from self ranks in a repertory grid (Ravenette, 1969). Children ranked themselves and seven other friends or relatives of similar age on eight provided constructs: the positive poles being, most likely to have friends, not upset their family, not upset their teachers, not be worried, enjoy school, be good at schoolwork, have interests and hobbies, the person I would most like to be. Constructs were selected following interviews with children from the unit. The rank order of the self element on each of the eight constructs was summed to give the second measure of self esteem. Scores could range from 8 to 64, with high scores indicating low self esteem.

(4) *Emotional and behavioural difficulties.* The Rutter B2 scale was completed by unit teachers for the unit sample and by form or year tutor for the matched group. Devised as a screening instrument (Rutter, 1967), the B2 gives scores which can be categorised into several areas of disorder; neurotic, antisocial, undifferentiated neurotic and antisocial, and no disorder.

RESULTS

The SPSS package (Nie *et al.*, 1975), was used for all statistical analyses. Table 1 lists the matrix of correlation coefficients and the descriptive statistics for the four variables, Coopersmith SEI, grid self ranks, verbal IQ and reading age.

TABLE 1

INTERCORRELATIONS BETWEEN SELF ESTEEM SCORES, VERBAL IQ, AND READING AGE, WITH DESCRIPTIVE STATISTICS FOR TOTAL SAMPLE

	1	2.	3.	4.	Mean	SD	Range
1. Coopersmith self esteem		0·53*	−0·03	−0·18	8·3	4·4	1-17
2. Total grid self ranks			−0·13	−0·37*	31·8	9·2	13-53
3. Verbal IQ				0·48*	89·1	11·4	69-125
4. Reading age					11·7	2·4	6·5-13·8

* $P<0{\cdot}01$.

The correlation between the Coopersmith scores and the grid measure of self esteem was statistically significant, with about 25 per cent of common variance between the two measures. This indicated a tendency for young people who selected items supposed to reflect low self esteem to also rank themselves towards the negative poles of the eight constructs. Correlations between these two measures were at a similar level for the subgroups of boys, girls and controls but not for the unit subgroup which had grid self rank scores of limited variance. Significant correlations between reading and verbal IQ for the total group and all subgroups were found as expected. However, there was no significant correlation between the SEI and verbal IQ, neither for the total group nor subgroups.

Correlations between verbal IQ and the other variables were also calculated for the 68 children who did the WISC-R verbal subtests. There were no significant differences in the levels of correlations compared with the total sample, 14 of whom did the WISC verbal subtests. There were no significant correlations between reading and the Coopersmith SEI for the total group, unit, control or female subgroups. For boys there was a significant low negative correlation ($r=-0{\cdot}24$, $P<0{\cdot}05$), indicating a weak tendency for higher reading levels to be associated with higher self esteem scores.

Analysis of variance showed significant self esteem differences between the Rutter B2 categories. ($F=5{\cdot}5$, $df=3, 75$, $P<0{\cdot}01$). The neurotic and undifferentiated groups had similar scores indicating lower self esteem than the antisocial and normal groups, as shown in Table 2.

TABLE 2

COOPERSMITH SELF ESTEEM SCORES FOR DIFFERENT CATEGORIES OF THE RUTTER B2 SCALE

	Mean	SD	N
Neurotic	11·9	3·4	12
Antisocial	7·7	4·0	19
Undifferentiated	12·0	4·4	3
No disorder	7·1	4·1	45
Total	8·2	4·3	79

Internal consistency of the 25 Coopersmith items was assessed with the Cronbach α formula, giving a coefficient of 0·77. Factor analysis of the 25 items using the principal factor direct method to obtain the factor matrix and the varimax solution for rotation, produced 10 factors with eigen values greater than unity. However, the first two factors accounted for only 17 per cent and 9 per cent of the total variance respectively. Items loading on factor one ($\pm 0{\cdot}30$) tended to relate to the subjects' perception of their emotional reactions, whereas those loading on factor two were a mixture of perceived popularity, self dissatisfaction and perceived problems with parents.

DISCUSSION

Contrary to prediction, self esteem as measured by the Lee short form of the Coopersmith SEI was not associated with general verbal abilities nor word reading. The one exception to this finding was a weak association between self esteem and reading for boys only. This suggests that boys may have been basing their self evaluations more on educational attainments than girls. One interpretation of the overall lack of association between the short form SEI and the educational indices is that none of the short form items refers directly to school activities. Even in the full SEI inventory only three of the 58 items refer to affective reactions to school attainment. This raises questions about the initial conceptualisation of self esteem and item selection, casting doubt on whether the short and full form inventories could be satisfactory measures of global self esteem when a basic area for self evaluation is not adequately sampled.

The association between the short form Coopersmith and the grid measures of self esteem indicates a limited degree of concurrent validity between the two measures of the same construct, only 25 per cent of common variance. This conclusion is supported by the finding that though reading was associated significantly with the grid measure of self esteem it was not associated significantly with the short form Coopersmith measure. Had reading been associated with both measures of self esteem, this would have suggested a stronger association between the two self esteem measures.

Despite the moderate α coefficient level, factor analytic findings indicated no general factor of self esteem. This is inconsistent with the Argyle and Lee results from which this short form inventory was designed This also counts against the validity of this inventory as a measure of a general self esteem construct. However, there was evidence that children who are judged by their teachers to have "neurotic" emotional and behavioural difficulties, e.g., often worries, tends to be fearful etc., had significantly lower Coopersmith self esteem scores than those categorised as antisocial or with no difficulties. There is some correspondence, therefore, not only between two self report indices of a self referent construct, but between a self report measure and judgments by others. This indicates that this short form Coopersmith inventory assesses some aspects of self perceptions and self evaluations about emotional reactions.

However, there is no evidence to conclude that what is assessed can be termed general self esteem. This raises doubts about the initial theoretical formulation of the self esteem construct, the programme of test construction and the procedure by which Lee developed a short form. It cannot be assumed that this version of the self esteem inventory is an adequate measure of a self esteem construct, despite the evidence of some internal and external relationships. More generally, this study illustrates the need for more theoretical clarification, and for adequate test construction and validation as Anderson (1981) has indicated. This would involve formulating a self esteem construct in specific theoretical terms as part of a wider theoretical framework. Such a formulation would be specific enough to identify several domains of self esteem and to list examples of the domain. Items could be written from these specifications and then subjected to various empirical analyses. Not only is this emphasis on content validity important in itself but it facilitates construct validity studies and therefore empirically based interpretation of the measure.

REFERENCES

ANDERSON, L. W. (1981). *Assessing Affective Characteristics in the Schools* Boston. Allyn and Bacon.

ARGYLE, M., LEE, V. (1972). *Social Relationships. Unit 10, E281.* Milton Keynes Open University Press.

BAGLEY, C., VERMA, C. K., MALLICK, K., and YOUNG, L. (1979) *Personality, Self Esteem and Prejudice.* London: Saxon House.

BURNS, R. B. (1979). *The Self Concept Theory Measurement, Development and Behaviour.* London· Longman.

CHAPMAN, J. W., and BOERSMA, F. J. (1980). *Affective Correlates of Learning Disabilities.* Lisse, Holland Swets and Zeitlinger.

COOPERSMITH, S (1967). *The Antecedents of Self Esteem.* San Francisco: Freeman

COOPERSMITH, S. (1975). Self concept, race and education. In VERMA, C. K., and BAGLEY, C. (Eds.), *Race and Education across Cultures.* London· Heinemann.

NIE, N., HULL, C., JENKINS, J., STEINBRENNER, K., and BENT, D. (1975). *Statistical Package for the Social Sciences*. New York: McGraw Hill.

PIERS, E V , and HARRIS, D. (1964). Age and other correlates of self concept in children. *J. educ. Psychol.*, 55, 91-95.

RAVENETTE, A. T. (1969). Method for administering the grid for children. Unpublished Manuscript.

ROSENBERG M. (1965). *Society and the Adolescent Self Image*. Princeton, N.J.: University Press.

RUTTER, M. (1967). A children's behaviour questionnaire for completion by teachers: preliminary findings. *J. Child. Psychol. Psychiat.*, 8, 1-11

SHAVELSON, R. J., BOLUS, R., and KEESLING, W. J. (1980). Self-concept recent developments in theory and methods *New Directions for Testing and Measurement*, 7, 25-43.

WATTS, A F. (1948). *Holborn Reading Scale*. London: Harrap

WELLS, L. E., and MARWELL, G. (1976). *Self Esteem, its Conceptualisation and Measurement*. London: Sage.

WYLIE, R C. (1961) *The Self Concept*. Lincoln: Nebraska University Press.

WYLIE, R. C. (1974) *The Self Concept*. Vol. 1. Lincoln· Nebraska University Press.

(*Manuscript received* 26*th July*, 1984)

Br. J. educ. Psychol., 55, 81-83, 1985

AN OBSERVATIONAL STUDY OF THE SOCIAL ADJUSTMENT OF SPINA BIFIDA CHILDREN IN INTEGRATED SETTINGS

BY LOW GUAT TIN
(*Institute of Education, Singapore*)

AND G. R. TEASDALE
(*The Flinders University of South Australia*)

SUMMARY Eight spina bifida children of average academic ability attending regular primary schools were matched with controls and systematically observed in the playground during recess and lunch breaks The disabled had significantly fewer interactions overall, interacted with fewer people, and spent more time alone or in dyads. Peers initiated significantly fewer interactions with the disabled, thus suggesting lower levels of acceptance. Underlying patterns of interaction also appeared to differ, especially in terms of setting and length. The disabled also interacted less with opposite sex peers.

INTRODUCTION

A prime rationale for the integration of disabled children in regular schools is the benefit supposedly accruing from social interaction with the non-disabled. It is argued that disabled persons who spend their lives in a society of 'normals' must learn to accept their differences and disabilities. Hence they must be socially prepared for life in the wider society, and a good way of doing this is to maximise their interaction with non-disabled people, especially during their childhood years.

An underlying assumption of this approach is that placing disabled and non-disabled children together for the entire school day will make social acceptance of each other inevitable. This assumption appears to be based on Homan's (1950) thesis that if the frequency of interaction between two or more persons increases then the degree of their liking for one another also will increase. Unfortunately, however, research on the social acceptance or rejection of disabled children in regular schools has been limited in scope, and outcomes have been inconclusive. There is certainly no clear support for the assumption of increased liking and acceptance.

Most research in this area has dealt with mentally retarded subjects, and results are confounded by variations in the degree of retardation, the nature of the integrated setting, and the training, experience and maturity of the class teacher. Typically, however, the retarded are less likely to be accepted, or are viewed neutrally or with indifference (Bruininks, 1978). There is evidence that retarded children in integrated settings are rejected more frequently than the physically disabled (Panda and Bartel, 1972). Nevertheless, in a sociometric study, Soldwedel and Terrell (1957) found a tendency for physically disabled children to be chosen less often by their non-disabled peers. Anderson (1973) likewise noted that a physical disability makes it more difficult for a child to establish effective social relationships. She found cases where there was so little contact between the physically disabled and others in the school that integration meant almost nothing. Allen *et al.* (1978), however, found that the four disabled children in their New Zealand study were socially accepted (according to teacher ratings and sociometric data) by the non-disabled. Playground observation of unsupervised activity also indicated little difference in patterns of social interaction.

In summary, it is clear that the presumed social benefits of integration require more careful investigation. Existing data on the social adjustment of physically disabled children in integrated settings are sparse, inconclusive, and based predominantly on indirect measures (sociometry, rating scales, etc.). It therefore was the purpose of the present study to supplement existing evidence using a more direct approach to the analysis of social interaction. The research design was based on the dual premises that school playgrounds provide the most effective laboratory for the study of children's social behaviour, and that systematic observation allows collection of the most authentic data.

METHOD

Subjects

All spina bifida children enrolled in years 5 to 7 in regular primary schools in the Adelaide metropolitan area were identified (N = 16). From this population were excluded children of below average academic ability (N = 4), and those whose parents or school principal did not grant permission for them to be observed (N = 4). The remaining children (one male, seven female) became subjects for a playground observation study, each being matched with a control from the same class on the basis of sex, age, ethnic background, SES and academic ability At the time of the study three subjects were in wheelchairs, two on crutches and three required no aids for mobility.

Procedure

Each subject and control was observed individually during two 15-minute recess breaks and two 25-minute lunch breaks. A minute by minute running account of all interactions was recorded. This was converted later to frequency counts and entered on a checklist based on predetermined coding definitions. A pilot study using a subject with rheumatoid arthritis and a non-disabled control provided the opportunity to practise observation procedures and refine coding definitions. Reliability checks with a second observer were carried out during the pilot study and on a periodic basis during actual data collection and coding.

Data analysis

Frequency counts were tallied for each coding category on an individual basis, thus allowing comparisons between subjects and controls using the Wilcoxon matched-pairs signed-ranks test.

RESULTS

Table 1 sets out the results of comparisons using the Wilcoxon test, and shows that the disabled had significantly fewer playground interactions overall than the controls, and

TABLE 1

SOCIAL INTERACTION IN THE PLAYGROUND. COMPARISONS BETWEEN SUBJECTS AND CONTROLS USING THE WILCOXON TEST (N = 8)

Coding categories	Sum of negative ranks*	Sum of positive ranks	Significance level
1 Total amount of interaction:			
(i) Total recorded (verbal, gestural, physical contact, facial expression, etc.)	32	4	<0·05
(ii) Total number of different people with whom child interacted	35	1	<0·02
2. No observable interaction for at least 30 seconds in any 1-minute period			
(i) Child quite alone	1·5	34·5	<0·02
(ii) Group play child in proximity but not interacting	24	12	NS
(iii) Group discussion: child in proximity but not interacting	2·5	33·5	<0·05
3. Size of group.			
(i) Child and 1 other	2	34	<0·02
(ii) Child and > 1 other	36	0	<0·01
4. Source of interaction·			
(i) Initiated by subject/control	14·5	21·5	NS
(ii) Peer initiated	36	0	<0·01
5. Setting of interaction			
(i) Play	36	0	<0·01
(ii) Discussion	4	32	<0·05
6. Length of interaction.			
(i) Long: i.e., > 30 seconds in any 1 minute	3	33	<0·05
(ii) Short: i e., < 15 seconds in any 1 minute	33	3	<0·05
7. Sex of peer with whom interaction took place.			
(i) Same sex	22	14	NS
(ii) Opposite sex	34·5	1·5	<0·01

* In all cases higher sums of negative ranks imply lower tallies for subjects, and vice versa.

interacted with significantly fewer people. As a corollary of this, the results also showed the disabled spending more time alone, and being less active participants in group discussion. Moreover the disabled interacted in dyads more frequently than the controls, and less frequently in larger groups. Although there was no significant difference in the extent to which subjects and controls initiated interaction with peers, it was clear that peers initiated significantly fewer interactions with the disabled.

Understandably there were significant differences in the settings in which interaction took place. For the disabled most interaction occurred during passive discussion, while for the controls it occurred during play. Consequently longer interactions were more frequent for the disabled, and shorter interactions for the controls. Finally, the disabled interacted significantly less frequently than controls with opposite sex peers.

DISCUSSION

A consistent pattern of social interaction has emerged from this observational study of spina bifida children in school playgrounds. Although each of the eight subjects had spent at least two years in integrated settings there was still evidence of social isolation: they had less interaction with peers and were more frequently alone or on the fringe of group processes. This aloneness was accentuated by the fact that peers initiated significantly fewer of the interactions that did take place.

Inevitably, of course, the more limited physical mobility of the disabled reduced their opportunity for interaction. Nevertheless, their basic pattern of interaction differed from that of the controls, thus suggesting a qualitatively different experience of social relationships The disabled interacted with fewer people for longer periods, typically in passive, one-to-one settings rather than in the give-and-take of larger, more active groups. They also has less contact socially with peers of the opposite sex.

These findings raise doubts about the presumed social benefits of integration. In this particular study physical proximity did not promote normal levels of playground interaction. The fact that peers initiated significantly fewer interactions suggests that the disabled had not achieved the same degree of social acceptance. And even when contacts were established there appeared to be differences in the underlying patterns of interaction, differences that resulted in reduced opportunities for the disabled to experience a normal range of social relationships.

ACKNOWLEDGMENTS. — Thanks are due to the Spina Bifida Association of South Australia, the Adelaide Children's Hospital, principals and teachers of co-operating schools, and Dr. B. Mageean of Flinders University.

REFERENCES

ALLEN, R , WILTON, K., and BALLARD, K. (1978). Social acceptance of physically handicapped children. Integration into regular primary schools *Aust. J. Special Educ.*, 2, 9-14.

ANDERSON, E. M. (1973). *The Disabled Schoolchild* London· Methuen.

BRUININKS, V L (1978) Peer status and personality characteristics of learning disabled and nondisabled students. *J. Learn. Disabilities*, 11, 484-489.

HOMAN, G. C (1950). *The Human Group.* New York. Harcourt, Brace and World.

PANDA, K. C , and BARTEL, N. R. (1972). Teacher perception of exceptional children. *J. Special Educ.*, 6, 261-266

SOLDWEDEL, B., and TERRELL, I (1957). Sociometric aspects of physically handicapped and non-handicapped children in the same elementary school. *Except Children*, 23, 371-372 and 381-382.

(*Manuscript received 5th May*, 1984)

Br. J. educ. Psychol., 55, 84-90, 1985

THE DEVELOPMENT OF YOUNG CHILDREN'S UNDERSTANDING OF THE PROCESS OF EVAPORATION

BY MICHAEL BEVERIDGE
(Department of Education, University of Manchester)

SUMMARY. This study looks at the development of young children's explanations of events which adults would explain as involving evaporation. Ninety children at each of 5, 7 and 9 years took part and these children were divided into three equal groups at each age A pilot study, which is also briefly reported here, had indicated that 7- and 8-year-old children gave explanations which were situation-specific. In particular, this suggested that a pan of water boiled away because the water "went into the pan". The main study went on to look at how different groups of children at each age responded to lessons about either the non-absorbent nature of the pan or the importance of the visible steam produced when the water boils. With age, the children were found to develop intuitive explanations of evaporation which became less inconsistent with the adult scientific view On the other hand there was only inconclusive evidence of the differential effects of the lessons at each age The failure of the lessons to immediately trigger off the evaporation as an explanation supports the importance of underlying cognitive structures.

INTRODUCTION

This study investigates the development of the concept of evaporation in young children. Their intuitive explanations of real world events involving evaporation are examined, and the effect of both providing evidence which contradicts their explanations and of directing their attention to relevant features of the situation are investigated. It is argued that this approach can give insight into the process through which scientific intuitions develop in children.

Piaget (1974) used a distillation apparatus to investigate children's concepts of evaporation. In his study the children were required to explain the passage of liquid from one part of the apparatus to another through a process of evaporation and condensation. Piaget found that 5-year-olds refused to admit any passage of matter had taken place. 7-8-year-olds admitted this transfer but without any idea of evaporation and condensation. 9-10-year-olds showed intuitions about the process but suggested that the liquid went into and out of parts of the apparatus; change of state was not used as an explanation. When children reached 11 years they used explanations which involved change of state. Piaget concludes that for the concept of evaporation to be formed the children require a corpuscular theory of matter, and 11-year-olds, in Piaget's study, did suggest that steam is tiny pieces of water.

Piaget suggests that it is the delay in forming this corpuscular theory which makes evaporation difficult to understand. When change of state does not require this theory to account for change of place children form correct intuitions more easily, e.g., 8-year-olds have no trouble with the ice/water transformation.

However, Piaget's study has two weaknesses as an indicator of the development of the evaporation concept. Firstly, the distillation apparatus is unlikely to have been part of the child's everyday experiences and they may well feel it to have a 'scientific' status to which their normal intuitions do not apply Secondly, in Piaget's study, the liquid does not 'disappear' completely; it changes its position. Explanations of 'passage of matter' are unlikely to be required in children's real world. In their experiences of evaporation they will see matter 'disappear' as in puddles drying up.

A more recent study of children's conceptions of the changes of state of water has been carried out by Osborne and Cosgrove (1983). This study revealed that although the majority (80 per cent) of 12-year-olds gave some form of 'gaseous' account of evaporation it was not until 16 years that 'bits of water in the air' became the form of this explanation. And Osborne and Cosgrove give no systematic evidence of the explanations offered by children younger than 12 years which would indicate how the 'gaseous' explanation develops.

Recent research has outlined a general framework for the development of scientific intuitions (Sutton, 1982; Driver, 1983; Gilbert and Watts, 1983). The implications are that immediate physical experience will be an important source of scientific intuitions, in conjunction with formal and informal instructions. This conclusion suggests that in the development of scientific concepts such as evaporation we need to chart the development of the relevant intuitions in typical contexts and under the influence of typical teaching about real world experience. The present study investigates the development of young children's concepts of evaporation by looking at how children's intuitive explanations of liquid evaporating are influenced by teaching which is directed in two different ways. The first approach was designed to show the children that a frequently cited explanation would not work. And the second lesson type attempted to draw their attention to relevant aspects of the process of evaporation.

METHOD

This study is in two parts: (1) a preliminary investigation into the general nature of young children's intuitions about evaporation and (2) a more systematic investigation of how such explanations given by 5-, 7- and 9-year-olds are influenced by either showing that their explanations conflict with experience, or emphasising relevant features of the situation which their intuitive explanations ignore and fail to account for.

PRELIMINARY INVESTIGATION

The preliminary part of the present study was in two parts. The first involved talking to children about events involving the process of evaporation to find out whether they did have any intuitive explanations and also to suggest what these might be. During this phase 40 7-year-old children were shown two pictures about each of the following events:

(*a*) puddles drying up in the sun
(*b*) clothes drying on radiators
(*c*) clothes drying on clothes lines
(*d*) ponds drying up.

The two pictures were, in each case, of before and after the drying had occurred. It should be emphasised that the pictures served as visual supports to the verbal descriptions of the events given by the experimenter. The children were questioned in an open-ended way to ensure maximum opportunity for their intuitions to be verbalised.

Two clear results were obtained from these interviews. Firstly, the process of evaporation, or any concept involving change of state of the liquid, was not correctly understood. Secondly, the children did have intuitive, and situation specific, explanations of the events presented. They had a strong tendency to describe clothes drying as being due to 'dripping' but the pond and puddle situations as the 'water goes into the ground'. The presence of the sun in these latter pictures 'makes the water go into the ground faster'.

In the second part of the preliminary investigation 36 8-year-old children were shown a pan of water being heated by a gas ring. When questioned as to what would happen if it were left on for a long time none of the children predicted that the water would 'disappear', and when shown the water evaporated completely in the pan they were extremely surprised. One comment given was "This does not happen at our house. There's still water in the pan because my mum makes the tea with it." Further questioning revealed a strong tendency for the children to give an 'absorption' explanation, e.g., "It's gone into the pan". These two preliminary investigations revealed that for each situation in which evaporation occurs there is a typical intuitive explanation being offered.

THE MAIN STUDY

The main study reported here examines the processes through which the concept of evaporation comes to replace these intuitions by examining the effects on children's explanations by giving them additional learning experiences of two types. The situation focused on is that previously described, water in a pan evaporating when heated by a gas flame. The typical response produced by the children in the preliminary studies implies lack of understanding both of the non-absorbent nature of the metal pan and of the connection between the steam rising off the boiling liquid and the reduction in its level. From these preliminary investigations it seemed possible that further instruction on these two aspects of the 'disappearance' of the water from the heated pan might assist the induction of an intuitive concept of evaporation.

A lesson explaining and illustrating the absorption properties of different substances should produce cognitive conflict with the "it goes into the pan" explanation, and a lesson which highlights the presence of steam in contexts of 'drying' should help children learn the association between evaporation and the change of state of the water.

Accordingly, the present study investigates the effect of these two types of instruction on children of three different ages.

The absorption lesson

The children were each presented with a small teat pipette and some water and were asked to put a few drops on to each of the following objects: a paper towel, a plastic saucer, a sponge, towelling, aluminium foil and a metal plate. The children were asked to carry on doing this until they were sure which objects the water would 'soak into'. They were then questioned about which objects the water soaked into, and then asked about the suitability of each material as a container for water. All the children learned to give correct answers.

Steam lesson

The children were shown an electric kettle containing water and asked "what happens when the kettle is turned on?" It was then turned on and attention was drawn to the steam. The children were encouraged further to describe other events in which they had seen steam. Situations such as bathrooms steaming up or liquid boiling were suggested if not spontaneously mentioned.

A mirror was placed in front of the boiling kettle and the condensation pointed out. And the windows of the room steamed up and this was discussed. Also a piece of wet cloth was passed around and then placed in front of an electric fire and dried completely. Attention was drawn to the 'steam'.

Ninety children at each of 5, 7 and 9 years took part in the study. Thirty children at each age received the steam lesson, 30 the absorption lesson and 30 'control' children received no formal instruction.

All 270 children were post-tested by showing them a circular aluminium pan of 10 cm diameter which they held and examined. The pan was then filled to ½ cm depth with water which was then heated over a Bunsen burner for six minutes, the time required for it to fully evaporate. Immediately following this the children were then asked "I want you to tell me about what has happened. I'll start off to make it a little easier for you. First I put the pan of water on the stove; now you tell me what happened next?" A further probe of "What happened to the water" was used if the children offered no explanation. The pan of boiling water was presented to the children in groups of six but the post-testing took place individually. Each group of six children contained two 'absorption', two 'steam' and two control children to ensure that minor differences in presentation would not bias group comparisons.

RESULTS

Mutually exclusive response categories were developed, organised around the absorption/steam and heat/no heat dimensions. Tables 1, 2 and 3 show the frequencies of the children's responses in each category at the three age levels. The responses were categorised independently by two persons with an overall agreement of 93 per cent. The responses about which disagreement occurred were then checked by the two observers together and agreement was reached with little difficulty.

Table 1 reveals that 5-year-old children seemed to learn little from either lesson. The two types of response which indicate that the pan was porous were given by 23 children, and answers which make reference to the steam by only four children. Also neither of the two types of lesson had any effect, as compared with the control group.

At this age, the most frequent types of explanation make reference only to the flame as the cause of the water's disappearance with no reference to the process involved. Rather than the lessons increasing the frequency of reference to the process it seemed that they had, if anything, the opposite effect.

Table 2 shows the 7-year-old children to be strongly inclined to refer to the process involved but to focus on the 'absorption' explanation. Furthermore, the lesson on absorption, rather than off-setting this tendency, seemed to encourage reference to 'soaking into the pan'.

Table 2 shows that the entirely transductive responses which only referred to heat do not occur in 7-year-olds. However, the steam lesson may have influenced the six children who did identify 'steam' as being associated with the disappearance of the water, but they appeared to view the steam as causal in a transductive manner, i.e., "the steam causes the water to disappear".

The 9-year-old children show a strong preference for the 'steam as a part of the process' type of explanation. 'Absorption' responses have greatly reduced in number by this age. The steam lesson does seem to be having a positive effect and the heat/steam connection is made by 23 of the 30 children who had the steam lesson. However, by 9 years of age, children who had not had the steam lesson tended, to a lesser extent, to suggest that steam had a part to play in the process. These references to steam did not, in general, indicate steam as a cause, as with 7-year-olds, e.g., "the steam makes it go away", but more as some puzzling aspect of the process which needed to be accounted for, e.g, "the flame makes the steam come and the water goes". However, explanations clearly involving the concept of change of state were not found at all.

In order to quantify the effects of both age and lesson type on the children's explanations two values of chi-square were obtained. The first calculation used the technique of Maxwell

TABLE 1

EXPLANATIONS GIVEN BY 5-YEAR-OLD CHILDREN FOR THE DISAPPEARANCE OF WATER IN A SAUCEPAN WHEN HEATED

Response	Group given previous lesson on absorption	Group given previous lesson on 'steam'	Comparison group with no lesson
Soaks into pan (−ref. to heat)	4	4	3
Goes into stove i.e., through pan	3	2	7
Steam in some way related	1	5	1
Heat the only aspect of the situation referred to, e.g , "the flame does it."	14	13	12
Don't know	8	6	7

TABLE 2

EXPLANATIONS GIVEN BY 7-YEAR-OLD CHILDREN FOR THE DISAPPEARANCE OF WATER IN A SAUCEPAN WHEN HEATED

Response	Group given previous lesson on absorption	Group given previous lesson on 'steam'	Comparison group with no lesson
Soaks into pan	20	14	13
Goes into stove i.e., through the pan	7	6	13
Steam in some way related, e.g , "its something to do with the steam."	0	6	0
Boils over	0	0	1
Don't know	3	4	3
Heat the only aspect of the situation referred to, e.g , "the flame does it."	0	0	0

TABLE 3

EXPLANATIONS GIVEN BY 9-YEAR-OLD CHILDREN FOR THE DISAPPEARANCE OF WATER IN A SAUCEPAN WHEN HEATED

Response	Group given previous lesson in absorption	Group given previous lesson on 'steam'	Comparison group with no lesson
Soaks into pan (+ or − ref. to heat)	6	1	6
Goes into stove i.e., through the pan	1	0	0
Steam in some way related, e.g., "something to do with steam."	10	6	7
Heat the only aspect referred to, e.g., "the flame does it."	0	0	3
Heat in association with 'steam'	13	23	14

(1961) for partitioning the degrees of freedom in contingency tables and was applied to the category of absorption responses. This technique divides the overall chi-square value into additive components which, in this case, are associated with either differences between the ages or differences within ages; the latter differences being those associated with different lesson types. The values of χ^2 obtained across the different ages was 46·7, df 2, revealing a highly significant effect of age on the production of absorption type responses. The within-age differences were, however, all non-significant: $\chi^2 = 0{\cdot}20$ for 5-year-olds, $\chi^2 = 3{\cdot}4$ for 7-year-olds and $\chi^2 = 2{\cdot}33$ for 9-year-olds.

This lack of statistical significance shows there was no clear effect of the differences between the two lesson types and the controls on absorption responses. However, the analysis takes no account of the direction of effect of the absorption lesson. The 20 absorption responses given by 7-year-olds after the absorption lesson does not produce statistical significance but does reveal a trend in the opposite direction from that which was expected. The lessons were designed to reduce, not increase, absorption responses. A second chi-square was carried out on the between-lesson type frequencies for the 9-year-olds' 'heat plus steam' responses which, it is argued, provide evidence of the development of an intuitive base for the scientific concept of evaporation. There was again no significant effect of lesson type ($\chi^2 = 3{\cdot}58$, df 2) although 30 per cent more responses in this category did follow the 'steam' lesson. The age effect on this type of response was again strong; no such responses were produced by the two younger groups of children.

DISCUSSION

The lessons given to the children were not lessons on evaporation. They gave the children experiences of either the absorption properties of materials or the presence of steam but did not put these phenomena together into a full explanation The purpose of the study was to see how the children used these experiences to build up an intuitive theory of evaporation for themselves. We find the 5-year-olds tending to produce phenomenistic explanations (Piaget, 1930). This association of two phenomena to produce the 'flame causes the water to disappear' may well be an example of transductive reasoning in which the child takes one perceptually compelling event and proceeds to relate it to another such event. This type of reasoning contains no appeal to logical necessity. And as Piaget noted, causal explanations by children at this age seldom include mediating links between cause and effect. The few 5-year-old children who gave an indication of the process, albeit incorrectly, made no reference to a cause.

The 25 per cent of 5-year-old children whose responses were categorised as 'don't know' are difficult to interpret. They may either mean "I can't think of anything" or "I can't think of anything which seems right". The latter response is the more cognitively advanced but the present study did not separate these two types of response.

The shift to process explanations by the age of 7 years, albeit to incorrect processes of absorption and permeability, is consistent with Piaget's (1930) finding that, of his 17 stages of development in understanding physical causality, the 10th stage, making a genuinely physical explanation, appears at 7-8 years old.

Despite the absorption lesson showing evidence that metal does not absorb water, this lesson seems to have led some of the 7-year-old children to an absorption explanation. Bruner (1966) and Gelman (1969) have found that taking children's attention from misleading attributes of conservation tasks can improve performance. In the present study the absorption lesson seems to have drawn some of the children's attention away from the relevant attribute of the steam and to have focused attention on the misleading absorption explanation.

The 9-year-old group showed a strong tendency to produce cause plus process explanations, which would not need radical revision for the principles of change of state from liquid to gas to be incorporated. But some 9-year-old children still favour the absorption explanation in both the control and absorption lesson groups, although permeability is only offered as a description of the process by one child. Hence even at the 9-year-old level there remains a range of intuitions which would be brought by the children to formal instruction about evaporation, and as a result some children will be required to do more cognitive work than others.

CONCLUSIONS

Several studies have shown that temporal contiguity is used by children to impute causality in novel situations (Siegler, 1975; Mandelson and Shultz, 1976). In the present study this tendency may assist the younger children to focus on heat as a cause, but the process itself does become intuitively important to the children at the relatively early age of 7 years.

Piaget (1930) suggested that this feeling for processes of nature stems from an interest in how machines work. Whether or not it is exposure to mechanistic explanations which is the prime influence, an increased number of process-oriented explanations seems to occur between 5 and 7 years.

As suggested earlier, noting the association of evaporation with steam does not mean the child understands the correct relationship. In this way causality can be inferred from contiguity, as in the suggestion that 'dark suits cause dandruff'. Or, in the present case, steam is a necessary causal factor for evaporation, and if no steam is seen some other process is involved.

Champagne and Klopfer (1981) suggest that directing children's attention to cognitive conflicts can be useful as a technique for developing scientific understanding. However, in the present study the presentation of apparently contrary evidence to children, while it did induce what appeared to be conflicting cognitions about the permeability of metal, did not lead the children to immediately resolve cognitive conflict into a more general and adequate theory of evaporation.

The recognition and resolution of conflicting cognitions may well be an important component in the development of scientific intuitions but the present study suggests that young children, without having contradictions pointed out to them, will hold apparently conflicting ideas without recognising the conflict.

The data described in the present study are consistent with Piaget's (1974) findings but also indicate the developmental processes which allow the scientific concept of evaporation to be understood. The children do have explanations of common experiences which develop so that a corpuscular theory can be incorporated to generate the concept of evaporation. The explanations offered by the corpuscular model are built on the processes of losing the absorption explanations and the development of the steam explanation. The corpuscular theory makes an explanation already accepted by the children more powerful but, as Osborne and Cosgrove (1983) found, for many children, despite considerable exposure to a science curriculum, the process of refining their explanations appropriately is carried on through their secondary school years.

ACKNOWLEDGMENT. — The author would like to acknowledge the assistance of Mr. William Senior in carrying out the pilot study

REFERENCES

BRUNER, J. S. (1966). On the conservation of liquids. In BRUNER, J., OLVER, R., and GREENFIELD, P. (Eds.), *Constraints on Learning*. London: Academic Press.

CHAMPAGNE, A. B., and KLOPFER (1981). Structuring process skills and the solution of verbal problems involving science concepts. *Sci. Educ.*, 65, 493-511.

DRIVER, R. (1983). Theories in action: some theoretical and empirical issues in the study of students' conceptual frameworks in science. *Studies in Science Education*, 10, 37-60.

GELMAN, R. (1969). Conservation acquisition. A problem of learning to attend to relevant attributes. *J. exper. child Psychol.*, 1, 144-162.

GILBERT, J. K., and WATTS, D. M. (1983). Concepts, misconceptions and alternative conceptions: changing perspectives in science education. *Studies in Science Education*, 10, 61-98.

MANDELSON, R., and SHULTZ, T. R. (1976). Covariation and temporal contiguity as principles of causal inference in young children. *J. exp. Child. Psychol.*, 2, 402-412.

MAXWELL, A. E. (1961). *Analysing Qualitative Data*. London: Spottiswoode.

NASS, M. L. (1956). The effect of three variables on children's conception of physical causality. *J. abnorm. soc. Psychol.*, 53, 191-196.

OSBORNE, R. J., and COSGROVE, M. (1983). Children's conceptions of the changes of state of water. *J. Res. in Science Teaching*, 20 (9), 825-838.

PIAGET, J. (1930). *The Child's Conception of Physical Causality*. London: Routledge and Kegan Paul.

PIAGET, J. (1974). *Understanding Causality*. New York. Norton.

SIEGLER, R. S. (1975). Determining the locus of developmental differences in children's causal reasoning. *J. exper. child Psychol.*, 20, 512-525.

SUTTON, C. (1982). Proceedings of conference on children's ideas about science. University of Leicester.

(*Manuscript received* 12*th September*, 1984)

Br. J. educ. Psychol., 55, 91-93, 1985

FURTHER VALIDATION OF THE BEHAVIOUR IN SCHOOL INVENTORY

By M. B. YOUNGMAN
(*School of Education, University of Nottingham*)

AND C. S. B. SZADAY
(*Student Services Branch, Education Department of Victoria, Australia*)

SUMMARY. Youngman's (1979) Behaviour in School Inventory is used as part of a counselling and guidance programme in an Australian secondary school. The sample comprises the complete year 7 intake (12- to 13-year-olds) of 96 boys and 37 girls. The results demonstrate good convergent and discriminant validity with correlations in the range 0·42 to 0 63 with the objective behavioural adjustment criterion, times sent out of class. It is argued that the BSI self-rating is particularly effective for girls, and generally superior to teacher ratings of adjustment

INTRODUCTION

The original validation of the Behaviour in School Inventory (BSI) appears in Youngman (1979). Briefly it is a self-report measure of behavioural adjustment to school designed to discriminate within the normal school population. This is an important qualification since several of the more commonly used behaviour measures are intended primarily to operate at the maladjusted end of the behavioural continuum. The self-report nature of the BSI is worth emphasising in view of disagreement over the validity of teacher ratings of behaviour. Dowling (1980, p. 26) contends that 'teacher or peer ratings of behaviour have been found to be the most effective predictors of future adjustment'. Conversely several studies have suggested that teacher ratings lack validity through contamination by perceived ability (Lambert, 1963), because of poor inter-rater reliability (McDermott, 1980), or because of differences in teachers' perceptions of boys and girls (Hartley, 1978). By incorporating Dowling's teacher rating, the present study aims to provide further validation of the BSI, and to compare its performance with a teacher rating of behaviour.

METHOD

Sample

The sample consists of the complete year 7 intake (12- to 13-year-olds) of a post-primary school in Victoria, Australia. Although the maximum sample size is 133 pupils (96 boys, 37 girls) Table 1 gives the exact numbers for each result. The boys and girls are analysed separately, partly because of the difference in their numbers, but primarily in response to earlier findings (Youngman, 1979) which demonstrated significantly higher BSI scores for girls.

Measures

Most of the variable details summarised in Table 1 are self-explanatory, but there are additional features of note. The teacher report ratings (variables 6-11) are composites from 11 subject areas assessed in the half-yearly reports, hence reliabilities are unknown. Variable 12 records the number of times during 1982 that the pupil was sent to a 'time-out' room as punishment for disruptive classroom behaviour, variable 13 gives the figure for the first term of 1983. Clearly these two variables have very high reliability, but there could be some concern over the distribution of a measure of extreme, and therefore possibly infrequent, behaviour. In the present sample the boys record a mean score of 6·6 (SD 7·9) and the girls 5·0 (SD 6·7). Whilst the relative sizes of the means and standard deviations do indicate some skewness (2 3 and 1 7 respectively) there is sufficient spread to permit valid interpretation of correlation coefficients. Since differences in the shapes of distributions of variables reduce correlations between them, the values reported here will be conservative estimates.

The teacher rating of pupil adjustment (variable 15; Dowling, 1980) has the original 4-point rating reversed so that here a high rating represents positive adjustment. The BSI subscales (variables 2-4) are studiousness, compliance and teacher contact respectively.

RESULTS

Table 1 reports product moment correlations listing the BSI and teacher rating variables in the columns, and the reference variables in the rows. Sample sizes and statistics for boys and girls are in separate rows.

The BSI overall adjustment score registers its highest profile of correlations with the two measures of disruptive behaviour, variables 12 and 13. Next in magnitude are the co-operation and effort ratings, with the lowest values for the performance and absence measures The subscale characteristics are similar to those of the total scale, differing in degree rather than kind. Separating the results for boys and girls has proved informative since the BSI correlations are consistently higher for the girls, in spite of their modest sample size.

The salient feature of the teacher rating results is the presence of low correlations with the BSI variables, and in consequence there is a different pattern of relationships with the reference variables. The highest values are recorded for the report ratings, not the disruption measures, and for the boys rather than the girls.

TABLE 1

PRODUCT MOMENT CORRELATIONS RELATING THE BSI AND TEACHER RATING MEASURES TO VARIOUS ADJUSTMENT AND PERFORMANCE REFERENCE MEASURES

		Adjustment variables				
	N	1	2	3	4	15
Reference variables	boys girls	BSI total	BSI study	BSI comply	BSI T-C	Teacher rating
5 Reading[a]	68	−08	−26*	04	10	20
	26	36*	38*	24	25	03
6 Achievement report June	94	25**	13	27**	17	49**
	37	54**	43**	55**	27	07
7 Achievement report Dec	89	32**	24*	30**	17	57**
	34	54**	43**	53**	30*	25
8 Effort report June	94	33**	20*	33**	20*	52**
	37	66**	48**	65**	47**	07
9 Effort report Dec	89	36**	25**	33**	24*	56**
	34	44**	31*	46**	25	27
10 Co-operation report June	94	32**	21*	31**	20*	46**
	37	73**	52**	71**	52**	00
11 Co-operation report Dec	89	41**	31**	39**	19*	56**
	34	51**	40**	55**	22	19
12 Times sent out 1982	96	−42**	−27**	−46**	−16	−49**
	37	−63**	−45**	−68**	−39**	−20
13 Times sent out 1983	89	−48**	−36**	−47**	−19*	−44**
	34	−53**	−41**	−56**	−23	−21
14 Days absent 1982	95	−10	−07	−10	−06	−25**
	36	−10	−05	02	−28	06
15 Teacher rating	95	23*	15	19*	18	
	35	−20	−28	−15	02	

Decimal points omitted from correlations

* P<0·05 ** P<0·01 (one-tailed test) [a] ACER (1973).

DISCUSSION

Convergent validity requires that the highest BSI correlations should be with the most relevant criterion variables, here the disruption measures. The order and size of the correlations clearly demonstrate strong convergent validity at a level of around 0·6 for girls and 0·5 for boys. Discriminant validity demands near zero correlations if the construct of interest is to be distinguished from similar but separate characteristics. Absenteeism offers the best discrimination test here and the correlations with that variable are effectively zero. Further support for discriminant validity arises from the correlation ordering which locates the achievement measures above absenteeism but below effort, co-operation and disruption. The teacher rating of adjustment does not fare so well when subjected to these two validity tests. Its highest correlations are with teacher reports rather than the disruption measures, indicating weak convergent validity. Discrimination also is not particularly well established. For boys the correlation with absenteeism is not zero, and for boys and girls correlations with achievement are not very different from those for adjustment.

It would appear from this pattern of relationships that at least two separate influences have diminished the validity of the teacher rating. Firstly the poor discriminant validity is consistent with Lambert's (1963) observations that teachers find it difficult to eliminate ability from ratings of other independent pupil characteristics. Secondly the inaccuracies in teachers' ratings of girls' adjustment could stem from the way teachers perceive the behaviour of boys and girls. Preconceptions about the rarity of assertive behaviour in girls (Hartley, 1978) could exaggerate their adjustment ratings, thereby reducing correlations with the disruptive behaviour criterion. Dowling (1980) himself concluded that there was little agreement between teachers' assessments and pupil self-ratings, and that the latter were better predictors of future adjustment. Clearly there remains considerable scope for clarifying the interrelationships between teacher ratings, pupil self-ratings, and actual behaviour. The BSI self-rating does seem to avoid some of the error associated with halo effects and sex bias in teacher ratings. The consistency of the findings of this modest validation study does support further use of the BSI as an effective measure for probing these relationships.

REFERENCES

AUSTRALIAN COUNCIL OF EDUCATIONAL RESEARCH (ACER) (1973). *Progressive Achievement Tests Reading Comprehension.* Hawthorn, Australia· ACER

DOWLING J. R (1980). Adjustment from primary to secondary school: a one year follow-up *Br J. educ. Psychol*, 50, 26-32

HARTLEY, D. (1978). Teachers' definitions of boys and girls. *Res in Educ*, 20, 23-35.

LAMBERT, P. (1963). The successful child· some implications of teacher stereotyping. *J. educ. Psychol.*, 56, 551-553

MCDERMOTT, P. A (1980). Sources of error in the psychoeducational diagnosis of children. *Psychology in the Schools*, 17, 12-24

YOUNGMAN, M B. (1979) Assessing behavioural adjustment to school. *Br. J. educ. Psychol.*, 49, 258-264

(*Manuscript received* 3rd *May*, 1984)

CONFERENCE ABSTRACTS

Br. J. educ. Psychol., 55, 94-107, 1985

Papers delivered at the Annual Conference of the Education Section of the British Psychological Society, held in the University of Cambridge, September, 1984.

Beech, J. R., and Harding, L. M., The New University of Ulster.

A DEVELOPMENTAL LAG PERSPECTIVE ON PHONEMIC PROCESSING AND READING RETARDATION

A link between the ability to process phonemes and reading ability has been indicated by many studies which have shown that a child who has difficulty in learning to read usually has difficulty in identifying the phonemic constituents of a word (Bradley and Bryant, 1978; Fox and Routh, 1980). It could be the case that children who are retarded in reading are behind because their phonemic processing capabilities are not yet developed (a developmental lag viewpoint) (Beech and Harding, 1984).

This experiment investigated whether there was such a developmental lag in skills associated with phonemic processing in backward readers, of normal non-verbal intelligence. Fifty-seven backward readers, 44 younger readers of the same reading level and 35 normal readers of the same age as the backward readers were given tests involving phonemic processing and memory. The control groups were both matched on non-verbal intelligence to the remedial readers. In virtually all these tests involving phonemic processing, including phonemic discrimination and speech tests, the backward readers were significantly worse than the chronological age controls, but there was no difference between the backward readers and the reading age controls. This finding that the younger readers were at exactly the same level of immaturity in phonemic processing as the remedial readers indicates a developmental lag at least within phonemic processes for the backward readers which is determining their potential level of reading performance.

The experiment also examined the relationship between reading style (whole words *versus* phonics) and phonemic processing There was little evidence for such a relationship nor were there differences in reading style between the three age groups.

REFERENCES

Beech, J. R., and Harding, L M (in press) Phonemic processing and the backward reader from a developmental lag viewpoint. *Reading Res Q.*, 19, 357-367.

Bradley, L., and Bryant, P. E. (1978). Difficulties in auditory organisation as a possible source of reading backwardness *Nature*, 271, 746-747.

Fox, B., and Routh, O. R (1980) Phonemic analysis and severe reading disability in children *J Psycholinguistic Research*, 9, 115-120

Beveridge, M., and Reid, D., Department of Education, University of Manchester.

DO PICTURES HELP CHILDREN UNDERSTAND WRITTEN INFORMATION?

Educational materials presented to children assume that pictures and diagrams help children understand written language. However, there has been little relevant systematic research on which to base this assumption. For example, the many studies of single work-picture associations shed little light on the ways normal children of secondary school age learn from books.

The psychological literature suggests a general picture facilitation of children's understanding of text. In this paper the authors discuss a number of their own studies which suggest that this "picture superiority effect" does not always occur. They examine some of the variables which influence the effectiveness of pictures on secondary school children's learning of text. Some of these studies consist of presentations of typical "textbook" pictures and text on microcomputers. The microcomputer-driven technique employed allows the individual strategies which children use in the learning process to be examined more directly than hitherto and the paper discusses some of the educational implications of these individual differences.

BIGGS, J. B., University of Newcastle, New South Wales, Australia.

THE DEVELOPMENT OF METALEARNING IN SECONDARY AND TERTIARY STUDENTS

In an earlier model of student learning, Biggs (1978) proposed three dimensions, each comprising a motive and a related strategy, to describe the domain of student learning, and suggested that learning was most effective when students' strategies were congruent with their motives for learning. The present paper reports a series of studies that has led to a revised model in which some theoretical assumptions, particularly those concerning congruence, are modified.

In view of the emphasis placed on metacognition in student learning (e.g., Entwistle, *et al.*, 1984) attention was paid in the present research to interactions between approaches to learning, personality and ability factors, situational factors, and the quality of performance. These interactions suggested a developmental model of *metalearning*, involving students' increasing awareness of their cognitive resources and of task demands, and the subsequent emergence of executive control over those resources to meet task demands. There appears to be a threshold of awareness, below which metalearning does not take place at all, and after which there is a gradual increase both in awareness and in control. Congruence between motives and strategies seems to be important mainly at a particular stage of metalearning development. Locus of control, information processing abilities, task specific knowledge, and certain general experiences interact to increase students' ability to approach learning wisely. The evidence for the revised model is reviewed, and some practical implications are suggested.

REFERENCES

BIGGS, J. B. (1978). Individual and group differences in study processes *Br. J. educ Psychol.*, 48, 266-279.

ENTWISTLE, N., HOUNSELL, D., and MARTON, F. (Eds.) (1984). *The Experience of Learning*. Edinburgh· Scottish Universities Press.

BILL, J. M., SYKES, D. H., and HOY, E. A., Departments of Education and Psychology, The Queen's University of Belfast.

PERFORMANCE OF VERY LOW BIRTH WEIGHT CHILDREN OF SCHOOL-AGE: A REVIEW

Due to modern neo-natal intensive care practices, many more very low birth weight (VLBW) infants now survive. Although they constitute, in general, only 1 per cent of all deliveries and while many appear (at least in terms of gross criteria) to develop normally, nevertheless with the wider spread of the most advanced intensive care techniques, it is expected that increasing numbers of high risk survivors will enter the school system, where they meet the first and most significant test of their capacity to function adequately in society. Evidence of performance at this stage takes on increased significance in the light of recent findings that assessments of degree and type of handicap made at early stages of development provide an uncertain basis for prediction of future status.

This review of the published literature of VLBW children of school-age examines the adequacy of the methodology and the appropriateness of the assessment techniques employed. It then attempts to extract conclusions as to the consistency of the findings and draws inferences for the future study and management of VLBW children.

CLAYFIELD, ROSEMARY G., and DAVIS, T N., University of London Institute of Education.

DEVELOPMENTAL CHANGES IN CHILDREN'S MODES OF REPRESENTATION USED BETWEEN 3 AND 9 YEARS

Ninety-six children aged 3 years, 5 years, and 7 years, drawn equally from upper and lower socio-economic groups and with equal numbers of boys and girls in each age group, formed the subjects of this study. They were tested on a version of the Oddity Problem (Lunzer, 1968) at seven to eight months intervals during a two-year period. The groups were thus aged 5 years, 7 years, and 9 years at the conclusion of the study. Four stages are identified in the administration of the task: the acquisition of a learning set, its generalisation to further one trial instances, the reproduction by the subject for the tester of an Oddity Problem using the task materials, and finally verbal justification of oddness as the concept attained. These four stages appear to represent levels of task difficulty; the ability to reflect on preceding actions and to give a verbal justification has been considered to be evidence of mental representational or operative ability (Inhelder and Piaget, 1964; Lunzer, 1968).

From a different standpoint, Bruner (1966) has suggested that there is a developmental order of emergence of modes of representation, enactive, iconic and symbolic, which once established are potentially interchangeable. However, apart from everyday observation, and inference from Piagetian studies, there is a dearth of direct empirical support for Bruner's hypothesis. This study investigates whether the four stages of oddity concept acquisition provide such evidence. The data categorised from the study are analysed in a variety of situations, by the McNemar test and the Cochran Q test, within population groups over time, within test situations, between socio-economic groups, and between sexes. Highly significant results indicate support for the theoretical position of Bruner (1966) that the enactive mode precedes symbolic representation through language as a means of representing experience. From 7 years, a marked increase in the ability to give a verbal justification is shown. The implications of these findings for classroom practice, and from the theoretical position, are discussed. Wider educational issues from the subgroups analyses are also considered.

REFERENCES

BRUNER, J (1966). *Towards a Theory of Instruction* Belknap Press

INHELDER, B., and PIAGET, J. (1964). *The Early Growth of Logic in the Child* London· Routledge and Kegan Paul.

LUNZER, E. A (1968) *The Regulation of Behaviour*. Staples Press.

CLAYTON, T., Specialist Senior Educational Psychologist, Wiltshire Psychological Service.

THE WORKSHOP APPROACH TO TRAINING TEACHERS IN BEHAVIOURAL SKILLS

Over the last seven years the Wiltshire Psychological Service has run a series of behavioural workshops for teachers. Typically, these have consisted of an introductory session during which the wider aspects of learning theory have been presented, followed by six two-hourly meetings devoted to identification of problem behaviours; observation, measurement and recording; selection of reinforcers; construction of programmes, classroom implementation of programmes; and evaluation. Usually they have catered for only a small number of participants, never more than 12, and have been staged by at least two educational psychologists. Teachers from ordinary primary and secondary schools have attended as well as those from special schools and units. For the most part the workshops have been general in that they have tackled a wide range of learning and behaviour difficulties but, more recently, some have focused on more specific themes such as Early Number (Clayton, 1982a). A Behaviour Modification Wallet containing lecture notes, example programmes, published articles, recording forms, practical exercises and a behavioural checklist has also been used (Clayton, 1982b).

All workshops and individual programmes were evaluated three months after completion and, more recently, some workshops have been evaluated on a longer term basis by questionnaire. Replies indicated that participants were most impressed by the practical and systematic approach of behavioural psychology and consequently rated the precise identification of problem behaviours and the systematic implementation of programmes as highly important workshop components. If our respondents are typical, it seems that the majority not only acquired skills which they themselves regarded as professionally useful but maintained these skills over a two- to three-year period and influenced colleagues and others in varying degrees.

REFERENCES

CLAYTON, W. T. (1982a). Early number workshop. a pilot in-service training project. *Remedial Educ.*, 17, 84-87.

CLAYTON, W. T. (1982b). The Behaviour Modification Workshop: an antidote to mindless technology *Behavioural Approaches with Children*, 6, 43-44.

COLMAR, SUSAN, Child Advisory and Psychological Services, Birmingham Education Authority.

BEHAVIOURAL INTERVENTIONS IN THE NATURAL LANGUAGE ENVIRONMENT: THEORY INTO PRACTICE

Language facilitation methods have traditionally tended towards two major emphases: the stimulation model favoured by more cognitively oriented educationalists and the behaviourally inclined structured training model (for example, DISTAR). It is argued that both of these approaches have disadvantages and a facilitation model is proposed which takes into account the research evidence on language learning in natural situations (Colmar and Wheldall, 1984).

Applied behaviour analysts such as Risley (1977) have stressed the natural environment as the situation where the facilitation of language learning should, and often does, take place. Instead of setting up special language rooms or indeed programmes, Risley emphasises the number of opportunities available for teaching language in everyday contexts. Qualitative features of the subsequent teaching-learning interactions can then be observed. It is possible to provide general and specific guidance to adult persons (whether parents, teachers or child care staff), in their typical situation, which will markedly alter the quantity and quality of their teaching-learning interactions with children. Thus a behavioural model of language teaching and learning in the natural settings used by young language learners (home, nursery, trips into the community) is proposed.

Preliminary findings from observational and experimental studies based on this model will be presented and their implications for practice discussed. These include courses for parents and teachers, a specially designed booklet for parents and an observation schedule for collecting data on naturally occurring language interaction sequences.

REFERENCES

RISLEY, T. (1977) The ecology of applied behaviour analysis. In ROGERS-WARREN, A , and WARREN, S. F. (Eds.), *Ecological Perspectives in Behaviour Analysis*. Baltimore University Press.

COLMAR, S., and WHELDALL, K (1984). Behavioural language teaching. using the natural language environment. *Behavioural Approaches with Children*, 8 (in press).

CONTI-RAMSDEN, GINA, Education Department, University of Manchester.

INTERVENTION STRATEGIES FOR LANGUAGE IMPAIRED CHILDREN IN ORDINARY SCHOOLS

The focus of the intervention strategies to be described in this presentation is on the child's development as an effective communicator. The term effective communicator entails that the child has learned the social utility of language. In order to understand this position it is necessary to review the historical factors that have given rise to it and the model of language learning on which it is based. This presentation, then, is divided into four areas: (*a*) a brief historical review of the content of child language intervention programmes; (*b*) a brief presentation of an interactive, unified model for language learning; (*c*) a discussion of the population of language impaired children, their definition and characteristics and (*d*) a discussion of the intervention strategies for promoting language in ordinary school settings.

DAVEY, A. G., Reader in Applied Social Studies, University of Newcastle upon Tyne.

RACIAL AWARENESS IN BRITISH PRIMARY SCHOOL CHILDREN

The aim of this study, financed by the Social Science Research Council, was to assess the effects of growing up in a multi-ethnic society on the ways in which children perceive themselves in relationship to others. The children, white, West Indian and Asian between the ages of 7 and 11 years, were drawn from 16 primary schools in the north and south of England. In each region, half the schools had more than 50 per cent West Indian or Asian children and half had 20 per cent or less. In addition, a small 'no contact' sample of white children was taken from rural schools where children had no experience of living in a multi-ethnic community.

A variety of measures were constructed to assess the children's perception of ethnic differences, the personal attributes they used as grouping criteria and their emotional response to people perceived as belonging to groups other than their own. A study was also made of the children's friendship patterns and all their parents were interviewed. The results provide clear evidence of the early onset of ethnocentricity. Ethnic distinctions are used extensively by primary school children in their attempts to make sense of the world. We found little desire for intergroup friendship amongst the children in any of the three age groups. Moreover, by the age of 7 or 8, they had not only learned their society's implicit system of classification but what attitudes should be adopted towards people classified in a particular way. The investigators concluded that, if intergroup acceptance is to be improved, collaborative group learning techniques need to be developed which increase the inter-dependence of children on each other and create a context for learning in which co-operative skills can be learned and practised.

REFERENCES

DAVEY, A. G., and MULLIN, P. (1980). Ethnic identification and preference of British primary school children. *J. child Psychol. Psychiat*, 21, 241-251.
DAVEY, A G., and MULLIN, P. (1982). Inter-ethnic friendship in British primary schools. *Educ Res.*, 24, 83-92.
DAVEY, A G. (1983). *Learning to be Prejudiced Growing up in Multi-ethnic Britain* London: Edward Arnold.

DAVIES, D. R., Psychology Department, Carlton Hayes Hospital, Leicester.

THE EFFECTS OF ROLE ENACTMENT ON CHILDREN'S PERFORMANCE ON TESTS OF CONVERGENT AND DIVERGENT COGNITIVE ABILITY

A study was carried out to investigate the effects of gender role enactment on children's performance. Forty-four children equally divided according to sex, aged 13 to 14 years and 16 to 17 years, were selected from the third and sixth forms of a Greater London comprehensive school.

Two tests were used to examine the effects on performance between a child's score on a test and the child's score on the test administered under the opposite gender role enactment condition. Two tests were employed differing on the dimensions of convergence-divergence and on the sex typing of tasks. The tests were a mechanical reasoning test (D.A.T.) and the uses of objects test (Hudson, 1968).

The results indicated that under the conventional administrative procedures boys scored better than girls on the mechanical reasoning test, whereas girls scored better than boys on the uses of objects test. However, the effect of opposite gender role enactment differed for the third and sixth formers. At the third form level boys' and girls' scores reflected differences in performance and on gender type dimension. At the sixth form level girls exhibited a slight but non-significant increase in mechanical reasoning scores and decrease in uses of objects scores, whereas boys' scores significantly decreased on both tasks. The results are discussed in terms of increased salience of gender-appropriateness of tasks at adolescence with reference to the results of a measure of gender role stereotyping B.S.R.I. (Benn, 1974), administered to the sample

REFERENCES

BENN, S. L. (1974). The measurement of psychological androngyny. *J. consult. clin. Psychol.*, 44, 155-162.
HARGREAVES, D. J. (1979). Sex-roles and creativity. In HARTNETT, D., BODEN, G., and FULLER, M. (Eds.), *Sex-role Stereotyping*. London: Tavistock.
HUDSON, L. (1968). *Frames of Mind*. Harmondsworth: Methuen.

ELITHORN, A., COOPER, R., and STAVROU, A., Institute of Neurology and the Royal Free Hospital, London.

VISUAL VERSUS AUDITORY IMAGE PREFERENCE AND THE CEREBRAL LATERALISATION OF VERBAL AND NON VERBAL SKILLS

Some recent studies have demonstrated that marked individual differences in some verbal and non-verbal skills can be related both to sex and handedness. In a study with young adults of graduate ability, left-handed men performed better than right-handed men on the Perceptual Maze Test. Left-handed women performed better on a lexical decision task than did the right-handed women (Elithorn *et al.*, 1983).

In the present paper we report on a similar study with children aged 8 to 13. In addition to the Perception Maze Test presented under both the with and without information conditions, these children completed a Lexical Decision Test, Reaction Time Tests to visual and to auditory stimuli with both the left and right hands. They also completed a brief questionnaire on the strength of their visual and auditory imaging skills.

A preliminary analysis of the tests so far undertaken with these children suggests that there is also a relationship between the strength of an individual's auditory images and the relative lateralisation of auditory skills as represented by his reaction times to simple auditory stimuli.

In the present paper the results of these pilot studies with both adults and children are briefly presented. The role that automated psychological tests might have in relating an individual's patterns of intellectual skills and his cognitive style to his or her performance at school are also discussed.

REFERENCE

ELITHORN, A., STAVROU, A., SAVAGE, I., and JAMES, I. (1983). Psychological testing, cerebral blood-flow and the laterality of intellectual functions. In MELDRUM, B. (Ed), *Cerebral Blood-Flow Metabolism and Epilepsy*. John Libbey

FARRELL, P., Department of Education, University of Manchester.

SELLING BEHAVIOURAL PSYCHOLOGY IN MENTAL HANDICAP—SOME LESSONS FROM THE EDY PROJECT

EDY has had a great impact across the whole of the United Kingdom. Since the materials were published by Manchester University Press in 1981, well over 3,000 copies have been sold and more orders are coming in all the time. Indeed the authors are about to embark on a second revised edition. The self-instructional format has enabled psychologists to train senior staff in schools as instructors and hence facilitate ongoing staff development and training. Consequently EDY would appear to have been an outstanding success in giving behavioural psychology away to teachers.

This presentation will examine this success in more detail. Firstly, the different criteria which can be used to evaluate the success of EDY will be examined in conjunction with the methodological issues involved. Secondly, a brief summary of the formal evaluative studies will be presented and thirdly, the results of a questionnaire to EDY users on factors which affect the long term implementation of EDY techniques will be presented and discussed (Farrell, 1982; Farrell and Sugden, 1984).

REFERENCES

FARRELL, P T. (1982). An evaluation of an EDY course in behaviour modification techniques for teachers and care staff in an ESN(S) school. *Special Education Forward Trends*, 9, 21-25.

FARRELL, P. T., and SUGDEN, M. (1984). An evaluation of an EDY course in behavioural techniques for classroom assistants in a school for children with severe learning difficulties. *Educ. Psychol.*, 4 (in press).

FRANCIS, HAZEL, Department of Educational and Child Psychology, University of London Institute of Education.

THE DEVELOPMENT OF READING

The concept of development has three faces, one looking towards natural or universal features of developing ability, another towards logical analysis of the structure of what is to be learned, and a third towards culturally imposed requirements and guidance.

Theories of developmental stages in the psychology of learning to read may have their roots in any of the three positions, but it is easy to confuse them.

In this paper an attempt will be made to use empirical studies and educational reports to show how developmental stages may be curriculum fostered, rather than logically or psychologically based in the process of learning to read.

REFERENCES

CHALL, J. (1938). *Stages of Reading Development.* New York: McGraw Hill.

DOWNING, J., and LEONG, C. K. (1982). *Psychology of Reading* New York: Macmillan.

FRANCIS, H. (1984) Reading development in school — recent research trends. In BENNETT, S. N., and DESFORGES, C. W. (Eds.), *Advances in Classroom Research* (Monograph of the British Journal of Educational Psychology). Edinburgh: Scottish Academic Press

GHUMAN, P. A. S., University College Wales, Aberystwyth.

A STUDY OF LOGICAL THINKING ABILITIES OF INDIAN CHILDREN USING INDIGENOUS MATERIAL

This research is part of an on-going project designed to explore systematically the thinking abilities of Indian children. Previous work in this field (Ghuman, 1975, 1980) has been comparative, but this study deliberately focuses on the logical thinking abilities of children in India. Approximately 150 children from two schools reflecting two distinctive social strata in Delhi were given a battery of tests. Two of the problems chosen were entirely of Indian origin: a kinship puzzle and a problem relating to the predicament of a holy man going for a pilgrimage.

Preliminary findings suggest that only a few were able to solve the kinship problem regardless of social background. This is a disappointing result in view of the perspective taken by researchers such as Cole and Bruner (1972) on cross-cultural research.

The rest of the data are being analysed on the other logical tests and findings presented at the Conference. Comparison with British children may be possible on some of the tests. Implications of the findings are explored for the education of ethnic minority children in the United Kingdom.

REFERENCES

COLE, M., and BRUNER, J. S. (1972). Some preliminaries to some theories of cultural difference. In *Yearbook of the National Society of the Study of Education.* University of Chicago.

GHUMAN, P. A. S. (1975) *The Cultural Context of Thinking a Comparative Study of Punjabi and English boys.* Slough: National Foundation for Education Research

GHUMAN, P A S. (1978) Nature of intellectual development of Punjabi children *Int J. Psychol.*, 13, 281-294.

KYRIACOU, CHRIS, Department of Education, University of York.

EXPLORING EFFECTIVE TEACHING USING RATING SCALES

Rating scales have been widely used to explore teaching, particularly by teacher educators in the assessment of student-teachers on teaching practice. Such scales come in a variety of forms, ranging both in number and specificity of the attributes listed (e.g., 'clarity of presentation', 'use of eye-contact', 'relationship with pupils') and in the type of rating scale employed (e.g., norm-referenced, criterion-referenced, analytic). A notable recent use of rating scales was that by HMI (1982) in the assessment of teaching by probationary teachers

This paper outlines the use of a set of rating scales developed by the author to explore differences between lessons given by a number of teachers in secondary schools, and considers whether such an approach can form a useful basis for research on effective teaching. A distinction is drawn between research which may be useful in identifying aspects of effective teaching which can contribute to initial training courses, and those aspects which may discriminate between established teachers in their relative effectiveness on specified criteria (e.g , on-taskness, pupil gain, improved attitudes) (see Kyriacou, 1983).

REFERENCES

HMI (1982). *The New Teacher in School.* London: HMSO

KYRIACOU, C. (1983). Two-factor theory of effective teaching *Education Section Review*, 7, 40-42.

LOWDEN, D. G., School of Education, University College of North Wales, Bangor.

ATTITUDES OF TEACHERS TO INTEGRATION

The research was carried out in an attempt to evaluate, amongst other things, the attitudes of teachers towards children who had been diagnosed, in pre-Warnock terms, 'Educationally Sub-normal (Moderate).'

Three main instruments were used: an original questionnaire, a Purdue Foundation Attitude Scale (shortened form) and a guided interview schedule. The first was used in two versions, Welsh and English, the second and third in English versions only. Personal observations were also taken into account where relevant.

An operational definition of 'integration' was used to give guidance to teachers (socially acceptable interaction with normal children in any aspect of the school day — mere physical proximity *per se* is not sufficient), but in the event, since most teachers had at least heard of the Warnock tri-partite definition, no one expressed difficulty over the use of the term.

This is not the place to discuss further the nature of the practice of separating one group of children from another: the practice exists. Teachers react to it and that reaction is the purpose of this paper.

MATHEWS, P. V., North Cheshire College.

RESEARCH, PRACTICE AND POLICY CHANGE IN MATHEMATICS TEACHING

This paper reports one aspect of an eight-month interactional investigation into the thoughts, feelings and actions in teaching-learning mathematics of two tutors, their 'learners' (17 practising teachers taking the Diploma in Mathematical Education and one second-year B.Ed.) and some of their pupils. It takes as its focus the formulation of a previously recognised, yet rarely acknowledged and inadequately resolved paradox, namely that we expect teachers to inspire confidence in pupils in order to act mathematically yet to act mathematically requires a spirit of doubt and puzzlement (Brown, 1971; Higginson, 1980). Hence the paradox that teachers should aim to inspire both confidence and doubt!

At the root of the paradox are two conflicting and incompatible teaching-learning perspectives. The immutabilist perspective with its emphasis on infallible authority, self-evident principles and immutable truths seeks to establish confidence through the acquisition of certain knowledge. This perspective manifests itself in facts-and-applications teaching-learning and research models, and is regarded as the major impediment to acting mathematically. The alternative mutabilist perspective provides a more adequate interpretation of the conditions for teaching-learning methematics. With its corresponding puzzles — paradoxes — problems — predicaments teaching-learning model, mutabilism emphasises change, communication, criticism and control. Individuals' T-L protocols will be used to illustrate the fortuitous manner in which the paradox was resolved through the identification of two interlocking sets of control conditions, best depicted on separate cusp-catastrophe surfaces (Woodcock and Davis, 1980).

It was recognised by all teachers concerned that the widespread enactment of the mutabilist-catastrophic T-L model poses a number of serious challenges to academics and professionals on the one hand and to administrators on the other. In the case of the former it compels us to reject (*a*) error-free views of teaching-learning, (*b*) the process/product dichotomy and (*c*) traditional developmental notions and task analysis perspectives. In the latter case it compels us to reappraise the nature of the relationships amongst policy making, research and practice.

REFERENCES

BROWN, S I. (1971). Rationality, irrationality and surprise. *Mathematics Teaching*, 55, 13-19.
HIGGINSON, W (1980) Ideas of the seventies. *Mathematics Teaching*, 90, 5-8.
WOODCOCK, A., and DAVIS, M. (1980). *Catastrophe Theory*. Harmondsworth: Penguin.

MERRETT, F., and WHELDALL, K., Centre for Child Study, Department of Educational Psychology, University of Birmingham

THE BEHAVIOURAL APPROACH TO TEACHING PACKAGE (BATPACK) II: EVOLUTION AND EVALUATION

This paper reports the on-going development of the Behavioural Approach to Teaching Package (BATPACK) (Wheldall and Merrett, 1984). Each successive version of BATPACK from Mark I to Mark III has required participants to evaluate the course from the points of view of length, structure, timing and usefulness whilst questions have been asked about the elements found to be more or less useful. Each new version has demonstrated a response to this formative evaluation. At the same time more objective, quantitative evaluations have been carried out. The effects of BATPACK training are expected to be demonstrable in terms of observable changes in teacher behaviour and pupil behaviour. BATPACK has been continually revised until it could meet this criterion. Subsequent experiments have demonstrated that BATPACK training achieves its objectives of changing teachers' classroom behaviour and that this leads to improvements in children's behaviour (e.g., Wheldall *et al.*, 1984).

Despite the success of BATPACK Mark III a further (final) revision was carried out in order to produce BATPACK Mark IV following consultation with BATPACK tutors who had run their own courses independently. The aim was to provide clearer and more detailed instructions for tutors, to re-order and to change the emphasis of some of the elements and to provide more and better video sequences of classroom behaviour for skills training.

Future plans include a follow-up module to BATPACK, focusing on the management of behaviour problems with particularly troublesome individual children, and a secondary school version of BATPACK.

REFERENCES

WHELDALL, K., MERRETT, F. E., and BORG, M. (1985). The behavioural approach to teaching package. an experimental evaluation. *Br. J. educ. Psychol.*, 55, 65-75.

WHELDALL, K., and MERRETT, F. E. (1984). Training teachers to be more positive: the behavioural approach. In BENNETT, S. N., and DESFORGES, C. W. (Eds.), *Recent Advances in Classroom Research.* (Monograph of the *British Journal of Educational Psychology.*) Edinburgh: Scottish Academic Press.

NEWTON, R. J., and MATHEWS, P. V., North Cheshire College, Warrington.

THE STUDY SKILLS PARADOX: "YOU CAN BRING A HORSE TO WATER BUT YOU CANNOT MAKE IT DRINK"

What prevents students from following apparently good sound advice? One might be forgiven for expecting students to be more than willing to adopt new and successful practices, if only to guarantee later success in examinations. Yet students rarely seem to follow even the best of advice. It is therefore tempting to hold students themselves responsible for any failure they may experience.

In this paper we would like to change the focus by asking what prevents teachers learning anything about teaching from the 'mistakes' of their students. A review of the literature and personal action research has led us to question the relevance of much of the advice offered to students and enabled us to offer a more appropriate procedural and conceptual teacher-learner framework for those teachers and students who wish to gain insight into, and control over, their own teaching/learning.

NEWTON, SARAH

USING THE CLASSROOMS OF ESN(S) ADOLESCENTS TO SOLVE 'REAL-LIFE' PROBLEMS

What do severely mentally handicapped adolescents do when something goes wrong for them in the classroom? Do they try to sort out the problems themselves, turn to someone else for help, or just wait and hope that their plight is noticed? What do their teachers think that they will do? Need the adolescents be observed in the situations, or can we rely on the teachers' predictions as a method of assessing each individual's ability to cope? These and other questions were addressed in a study of problem solving by 21 ESN(S) adolescents (mean C.A. 15:10 yrs) and 31 non-handicapped children (mean C.A. 6:5 yrs) (Newton, 1981). This paper looks at the responses to the two problems that were presented to all of the class-members, unsharpened pencil and broken scissors.

The responses of the adolescents and children to the pencil and scissors problems were similar; they both tended to obtain solutions independently in the former situation, and to seek assistance with the scissors problem. The adolescents made more attempts to resolve the problems but they showed greater independence in obtaining solutions. The teachers in both groups accurately predicted who would resolve the problems, but were less successful in predicting the methods they would use to effect solutions. Where incorrect predictions were made, the teachers tended to underestimate the abilities of their pupils.

The study generated further questions concerning the differences between the ESN(S) and non-handicapped classrooms as settings conducive to the development of problem-solving skills, and the relationship between the teachers and their pupils with reference to the influence the former have on the problem-solving abilities expressed by the latter. Suggestions were made as to how problem solving skills could be taught without encroaching to any great extent on the teachers' already busy timetables (e.g., incidental teaching, Hart and Risley, 1980).

REFERENCES

HART, B., and RISLEY, T R. (1980) In vivo language intervention: Unanticipated general effects. *J. appl. Behav. Anal.*, 13, 407-432.

NEWTON, S. (1981) Social aspects of communicative competence in mentally handicapped adolescents and non-handicapped children. Unpublished doctoral dissertation, University of Manchester.

NORWICH, BRAHM.

ASSESSING PERCEIVED SELF EFFICACY IN RELATION TO MATHEMATICS TASKS: A STUDY OF THE RELIABILITY AND VALIDITY OF ASSESSMENT AND THE ACCURACY OF SELF EFFICACY JUDGEMENTS IN RELATION TO PERFORMANCE LEVELS

Though research into perceived self efficacy (Bandura, 1977) has originated in the clinical field, there has been growing interest in applying self efficacy ideas to the field of educational achievement (Schunk, 1981, 1982). This paper reports the results of a two-stage study of certain reliability and validity questions arising out of the use of two modes of self efficacy assessment — a direct and a generalised mode — in relation to particular classroom mathematics tasks for 9-10-year-old boys and girls in four schools.

The results for the first stage established the test/re-test, internal reliability and criterion validity of two measures used in the second stage — self-perception of maths ability, and a teacher rating of helpless behaviour in mathematics. In the second stage, perceived self efficacy was assessed in a repeated measures design over five trials in relation to particular maths tasks. Performance at the tasks was assessed after the first and fourth self efficacy assessment. Analysis of the results showed moderate to high retest and inter-interviewer reliability for the self efficacy measures. No significant differences between the direct and generalised modes of assessment were found. Results also indicated an increase in accurate self efficacy judgments after the first performance. Regression analysis indicated that children higher in self perception of maths ability and those lower in helpless behaviour in maths tended to make higher self efficacy judgments. The results were also analysed in terms of the relative contribution of the affective variables and prior task performance as predictors of subsequent self efficacy levels. The paper concludes with a discussion of the implications of the findings for future research in the field.

REFERENCES

BANDURA, A. (1977). Self efficacy: toward a unifying theory of human behaviour. *Psychol. Rev*, 84, 191-125.

SCHUNK, D. H. (1981) Modelling and attributional effects on children's achievement: a self efficacy analysis. *J. educ. Psychol.*, 73, 93-105

SCHUNK, D H (1982). Effects of effort attributional feedback on children's perceived self efficacy and achievement. *J. educ. Psychol*, 74, 548-556.

PUMFREY, P. D., Department of Education, University of Manchester.

THE BRADFORD 'BOOK FLOOD' EXPERIMENT: LESSONS FOR TEACHERS AND RESEARCH WORKERS

Children's reading attainments, attitudes to reading and reading habits are central educational concerns. A three-year longitudinal study compared the progress on these variables of 3,234 pupils initially aged 10 years attending four middle schools in Bradford. Two schools (deemed experimental) each received approximately 4,500 books in addition to their normal stock and capitation allowance for books. The two matched control schools received no such additional supply of books. It is from the viewpoint of consultant to the research that this paper is written (Pumfrey, 1983).

No simple *differential* gain consequent on the extra supply of books was identified in relation to either reading attainments or attitudes to reading, although important and statistically significant changes in the former occurred. The data obtained concerned the number and variety of books read by the pupils attending the four schools over the three years of the research represent a unique record of the development of children's reading habits. It provides important insights into the development of the reading habits of pupils attending the four schools involved in the study. In addition, case studies of avid and infrequent readers, together with analyses of the policies and practices of the four schools concerning the use of books and the teaching of reading, were carried out.

In order to highlight practical and methodological lessons of interest to teachers and research workers, the results obtained and the methodologies employed are considered. Specific criticisms of the design of the study, the analyses, resources and the research report are presented. It is concluded that the 'Book Flood' experiment had many shortcomings from the point of view of an empirical psychometric research orientation but that these are offset by the complementary data obtained using a thematic analysis.

REFERENCE

PUMFREY, P. D. (1983). Some effects of the Bradford 'Book Flood' experiment on children's reading attainments, attitudes to reading and reading habits. *Division of Educational and Child Psychology Occasional Papers*, 7, 3-7.

RICE, V., BOYLE, T., and TATUM, S., Liverpool Polytechnic.

I MAY BE YOU BUT VICI IS ALWAYS VICI: PRAGMATIC DEFICITS AND COMPENSATORY TACTICS IN AUTISTIC SPEECH

A number of studies of the speech of autistic children have reported marked pragmatic deficits (Baltaxe, 1977; Langdell, 1980). The aim of the present study was to explore this area in more detail by setting up a structured situation, a simple game, which permitted pragmatic demands to be systematically varied. The five subjects in the study (aged 8:9 to 16:7) attended a special school for autistic children. The phases of the game varied the pragmatic demands on the children in terms of such factors as person deixis (use of "I" and "you") and taking the perspective of the listener into account. The children's utterances in each phase were coded for functional appropriateness and structural elaboration.

The results indicated that as the pragmatic loading of the conditions increased the percentage of appropriate utterances fell, in certain cases very sharply. Despite these problems the children invariably succeeded in producing a correct response in the end. Examination of how they achieved these responses revealed the use of a series of distinct 'tactics'. A successful conclusion to the interactions was thus nearly always reached even though the adult and subject seemed to be operating from different ground rules. Many of the features of autistic speech regarded as bizarre were found in the study apparently to have a functional significance for the children. The main implication of the study is thus the need to be more sensitive to the autistic child's use of language from the child's point of view.

REFERENCES

BALTAXE, C. A. M. (1977). Pragmatic deficits in the language of autistic adolescents. *J. Paediatric Psychol.*, 2, 176-180.

LANGDELL, T. (1980). Pragmatic aspects of autism or why 'I' is a normal word? Paper presented at the Annual Conference of the Developmental Section of the British Psychological Society, Edinburgh, September, 1980.

RIEM, R., and DURKIN, K., University of Kent at Canterbury.

ONE-ONE CORRESPONDENCE IN COUNTING: WHAT MOTHERS CAN TEACH TEACHERS

Since Piaget, most if not all psychological models of counting development have been dominated by the notion of one-one correspondence. It has been treated as an independent and essential subcomponent of the mature skill which may be taught and tested without the context of counting activity itself (Piaget, 1952, Potter and Levy, 1968). This idea has been taken up by educationalists; Hargis (1971) has suggested that teaching retarded children to understand grammatical forms associated with one-one correspondence will help them learn to count.

By the above perspective, an intensive longitudinal study of how counting develops within five preschool mother-child dyads has produced some surprising results. Mothers treat one-one correspondence not as a logical component to be 'understood' by the child, but as a feature of counting activity which will be grasped incidentally as the child learns to count (*a*) in increasingly different task situations, and (*b*) with increasingly less other-regulation. Consequently, children learn about one-one correspondence with little or no formal teaching.

Apart from giving teachers valuable insight into what is attained before formal schooling, this paper explores the possibility of exploiting preschool strategies that emerge spontaneously in mother-child interaction.

REFERENCES

HARGIS, C. H. (1971) The significance of the grammar of one-to-one correspondence in teaching counting to the retarded. *Education and Training of the Mentally Retarded*, 6, 170-171.
POTTER, M C., and LEVY, E. I. (1968). Spatial enumeration without counting. *Child Dev.*, 39, 265-273.

SOLITY, J., Psychological Services, Walsall Education Authority.

DIRECT INSTRUCTION: IMPLICATIONS FOR TEACHING CHILDREN WITH SPECIAL NEEDS IN THE ORDINARY SCHOOL

The paper considers the following areas:

(1) Applications of behavioural psychology to teaching children with special needs. Advantages and limitations.
(2) Principles of Direct Instruction.
(3) Applications of D.I. to the teaching of literacy and numeracy skills.
(4) The implications of D.I. to teaching children with special needs.

In recent years psychologists and teachers have become increasingly interested in applying behavioural psychology to teaching pupils who have presented as having learning difficulties. Task analysis, teaching to behavioural objectives and Precision Teaching, which all have their roots in behaviourism, have been extensively developed with considerable success. However, concern has often been expressed that the approaches:

(*a*) do not provide a basis for teaching children to generalise skills,
(*b*) do not provide clear guidelines for deriving a curriculum,
(*c*) over-simplify the teaching of reading and arithmetic.

Direct Instruction goes some way to meeting these criticisms by trying to answer the question, "Why is the most efficient way to TEACH literacy and numeracy skills?" rather than the more frequently posed "How do children typically LEARN literacy and numeracy skills?" This is achieved by designing curricula in such a way which teach children skills which can then be generalised. Teaching generalisations depends on careful analysis of the concepts children are to be taught and identifying critical samenesses and critical differences between concepts.

The principles of Direct Instruction lead to clear guidelines being formulated on how best to prepare curricula for teaching reading and arithmetic skills. Furthermore Direct Instruction also provides a framework for organising instruction and devising presentation techniques.

Finally the theory of Direct Instruction has clear implications for both teachers and psychologists when completing assessments and preparing advice for LEAs under the terms of the 1981 Education Act. The issues raised concern the extent to which the assessment should focus on instructional variables rather than child variables.

STURMEY, P., Department of Psychology, University of Liverpool.

PORTAGE GUIDE TO EARLY EDUCATION: A REVIEW

The Portage Guide to Early Education has become a popular form of early intervention with mentally handicapped pre-schoolers. This review critically evaluates the service package consisting of a pyramidal management structure, a set of teaching materials and parents teaching their child with help from weekly home visits, and the outcome studies of Portage. Recommendations are made for future research and practice. Specifically the areas of training parents in general behavioural principles, modifying the written materials, and investigating mediating variables such as family and home characteristics are suggested as areas for future research.

SYKES, D. H., BILL, J. M., and HOY, E. A., Departments of Education and Psychology, The Queen's University of Belfast.

DEVELOPMENTAL PSYCHOPATHOLOGY AND THE VERY LOW BIRTH WEIGHT INFANT

Within the school-age population a significant minority of children fail to achieve their optimal academic and social effectiveness. They are characterised by school failure, poor peer relations, and a broad spectrum of anti-social behaviours. Such children, in addition to being the subject of concern within the educational setting, are believed to contribute disproportionately to the pool of adult psychopathology (Kohlberg *et al.*, 1972; Robins, 1966, 1978).

Contemporary approaches to the study of such children emphasise the developmental perspective. Developmental psychopathology has been defined by Sroufe and Rutter (1984) as "... the study of the origins and course of individual patterns of behavioural maladaptation..." Within this perspective, the study of high risk children is paradigmatic.

This paper looks at the very low birth weight (VLBW) child as a child at risk. Such children, it is argued, being relatively small in number, can serve as a lens through which the pattern of behavioural adaptation/maladaptation can be more accurately observed. The paper reviews the literature on the characteristics of VLBW infants that, together with significant environmental and social variables, determine the affective, social and cognitive adjustment/maladjustment of these children.

REFERENCES

KOHLBERG, L., LACROSSE, J., and RICKS, D. (1972). The predictability of adult mental health from childhood behaviour. In WOLMAN, B (Ed.), *Manual of Child Psychopathology*. New York. McGraw-Hill.

ROBINS, L. (1966). *Deviant Children Grown Up*. Baltimore: Williams & Wilkins.

ROBINS, L. (1978). Sturdy childhood predictors of adult antisocial behaviour: replications from longitudinal studies. *Psychol. Medicine*, 8, 611-622.

SROUFE, L. A., and RUTTER, M. (1984). The domain of developmental psychopathology. *Child Dev.*, 55, 17-29.

WHELDALL, K., and MERRETT, F., Centre for Child Study, Department of Educational Psychology, University of Birmingham.

THE BEHAVIOURAL APPROACH TO TEACHING PACKAGE (BATPACK) I: BACKGROUND

The Behavioural Approach to Teaching Package or BATPACK was developed by a project team based in the Centre for Child Study at the University of Birmingham. The aim was to develop a package which would train practising primary and middle school teachers in a number of key behavioural classroom management skills. BATPACK was informed by the programme

of experimental research into the behavioural approach to teaching carried out by Wheldall and Merrett and associated students in the Centre for Child Study (Wheldall and Merrett, 1984).

It was also informed by the attempts of other researchers to develop similar packages (Merrett and Wheldall, 1984). A brief review is given of these attempts and the way in which lessons learned from these were instrumental in developing the package into its present form.

BATPACK is designed to be taught by a tutor who has attended a special training course and who has a good working knowledge of behavioural psychology. The BATPACK course consists of six one-hour sessions called units, taught at weekly intervals. For each unit there are five objectives. Every unit is divided into five elements so that each corresponds to an objective. Elements are used to explore practical problems (identifying problem behaviours), to develop skills (pin-pointing, praising, observing) and to explain techniques (setting rules, defining work demands). Course members are also asked to observe and record their own behaviour. Every unit has an accompanying 'Takehome' in which an attempt is made to supply some of the theoretical material which will inform the practical skills learned in the unit and which provides the reading assignment for the week (Wheldall and Merrett, 1984).

REFERENCES

MERRETT, F. E., and WHELDALL, K. (1984). Training teachers to use the behavioural approach to classroom management: a review *Educ. Psychol.* (in press).

WHELDALL, K., and MERRETT, F. E (1984). The behavioural approach to classroom management. In FONTANA, D (Ed.), *Behaviourism and Learning in Education* (Monograph of the *British Journal of Educational Psychology*.) Edinburgh: Scottish Academic Press

WHELDALL, K., and MERRETT, F. E. (1984). Training teachers to be more positive. the behavioural approach. In BENNETT, S. N., and DESFORGES, C. W. (Eds.), *Recent Advances in Classroom Research.* (Monograph of the *British Journal of Educational Psychology*.) Edinburgh Scottish Academic Press.

WIGLEY, VERONICA, Hester Adrian Research Centre, The University of Manchester.

IMPLEMENTING BEHAVIOURAL PROGRAMMES: AN EVALUATION OF A PRIMARY SCHOOL PROJECT

This paper describes the evaluation of the work of Teacher-Child Interaction Project, a collaborative project between staff of the University of London Institutes of Education and Psychiatry and the ILEA (Berger, 1982; Yule *et al.*, 1984). The major goal of the project was to foster positive teacher-child interaction as a means of reducing the mutual difficulties experienced by teachers and a proportion of children in the classroom. It aimed to achieve the goal through the application of principles and techniques of behaviour modification adapted to the special needs and circumstances of teachers and children in local schools.

After pilot work a control group study was carried out to investigate the effects of the work of the project on teachers' and children's behaviour in the classroom. Observational data were collected by observers and by the teachers themselves and questionnaires were used to obtain information on teachers' opinions about the relevance and value of the work. The findings of this study are described, focusing particularly on the changes in the teachers' and children's behaviour in the classroom as monitored by observers. These findings are discussed with particular reference to variables which appear to be critical to the implementation of innovative strategies.

REFERENCES

BERGER, M (1982). Applied behaviour analysis in education a critical assessment and some implications for training teachers. *Educ. Psychol.*, 2, 289-300.

YULE, B, BERGER, M., and WIGLEY, V. (1984). Behaviour modification and classroom management In FRUDE, N., and GAULT, H (Eds.), *Children's Aggression at School.* Chichester: Wiley

PUBLICATIONS RECEIVED

Books for review have been selected from the following
('n.p.' means 'no price mentioned').

ADAMS, R. J. (1984). *Sex Bias in ASAT?* Victoria: ACER, pp. 154, $10·95.

AINLEY, J., BATTEN, M., and MILLER, H. (1984). *Staying at High School in Victoria.* Victoria: ACER, pp. 151, n.p.

ANDERSON, R. C., OSBORN, J., and TIERNEY, R. (1984). *Learning to Read in American Schools: Basal Readers and Content Texts.* London: Lawrence Erlbaum Associates, pp. 308, £24·50.

BEAUGRANDE, R. de (1984). *Text Production: Toward a Science of Composition. Volume XI.* New Jersey: Ablex, pp. 399, n.p.

BEE, H. L., and MITCHELL, S. K. (1984). *The Developing Person: A Life-Span Approach.* New York: Harper and Row, pp. 653, n.p.

BEECH, J. (1984). *Learning to Read.* Beckenham: Croom Helm, pp. 147, £13·95 (h/b), £7·95 (p/b).

BOWLBY, J. (1984). *Attachment and Loss: Volume I Attachment.* London: Penguin, pp. 425, £5·95.

BRETHERTON, I. (Ed.) (1984). *Symbolic Play, The Development of Social Understanding.* Florida: Academic Press, pp. 375, £34·50.

BROWN, R. I. (1984). *Integrated Programmes for Handicapped Adolescents and Adults.* Beckenham: Croom Helm, pp. 259, £14·95.

CANTOR, Dr. S. (1984). *The Schizophrenic Child.* Milton Keynes: Open University Press, pp 129, £5·95.

CEDEFOP, (1984). *Vocational Training Systems in the Member States of the European Community. Comparative Study.* Berlin: CEDEFOP, pp. 479, £6·90.

CROALL, J. (1984). *Neill of Summerhill.* London: Ark Paperbacks, pp. 436, £3·95.

CRONBACH, L. J. (1984). *Essentials of Psychological Testing* (*4th edit.*) London: Harper and Row, pp. 630, £11·95.

DAVIS, R. B. (1984). *Learning Mathematics: The Cognitive Science Approach to Mathematics Education.* Beckenham: Croom Helm, pp. 392, £22·50.

DELAMONT, S. (Ed.) (1984). *Readings on Interaction in the Classroom.* London: Methuen, pp. 350, £5·95.

DENT, N. J. H. (1984). *Moral Psychology of the Virtues.* Cambridge: Cambridge University Press, pp. 223, £20·00.

DUNHAM, J. (1984). *Stress in Teaching.* Beckenham: Croom Helm, pp. 176, £12·95.

DUNN, J. (1984). *Sisters and Brothers.* London: Fontana, pp. 155, £2·95.

EGAN, K. (1984). *Education and Psychology: Plato, Piaget and Scientific Psychology.* London: Methuen, pp. 210, £4·95.

FONTANA, D. (Ed.) (1984). *The Education of the Young Child.* Oxford: Basil Blackwell, pp. 315, £7·50 (p/b), £22·50 (h/b).

FOX, B. (1984). *The Law of Special Educational Needs.* Beckenham: Croom Helm, pp. 134, £12·95.

FRANCIS, H. (1984). *Minds of Their Own. An Inaugural Lecture.* London: University of London, Institute of Education, pp. 27, £1·50

FROEHLICH, W. D. (Ed.) (1984). *Psychological Processes in Cognition and Personality* New York: Hemisphere, pp. 284, $47·95.

FRUDE, N., and GAULT, H. (Eds.) (1984). *Disruptive Behaviour in Schools.* Bognor Regis: John Wiley and Son, pp. 234, £16·95.

GILBERT, R. (1984). *The Impotent Image: Reflections of Ideology in the Secondary Curriculum.* London: Taylor and Francis, £14·25 (h/b), £8·95 (p/b).

GREENE, J., and HICKS, C. (1984). *Basic Cognitive Processes.* Milton Keynes: Open University Press, pp. 128, £5·95.

GOOD, T. L., and BROPHY, J. E. (1984). *Looking in Classrooms.* New York: Harper and Row, pp. 424, n.p.

HALPERN, D. F. (1984). *Thought and Knowledge: An Introduction to Critical Thinking.* London: LEA, pp. 402, £34·95 (cloth), £17·50 (paper).

HARRIS, J. (1984). *Spoken Irish in Primary Schools.* Dublin, pp. 153.

HEPPLE, B. A. (1984). *Equal Pay and the Industrial Tribunals.* London: Sweet and Maxwell, pp. 66.

HOLLEY, C. D. and DANSEREAU, D. F. (Eds.) (1984). *Spatial Learning Strategies, Techniques, Applications, and Related Issues.* Florida: Academic Press, pp. 328, $45·00/£31·50.

Holt, J. (1984). *How Children Learn.* Harmondsworth: Pelican Books, pp. 303, £2·95.

Howe, M. J. A. (1984). *A Teacher's Guide to the Psychology of Learning.* Oxford: Basil Blackwell, pp. 162, £4·95.

Jeeves, M. A., and Greer, G. B. (1984). *Analysis of Structural Learning.* Florida: Academic Press, pp. 269, £25·00.

Jeffree, D., and Cheseldine, S. (1984). *Let's Join In.* London: Souvenir Press, pp. 135, £6·96 (h/b), £4·95 (p/b).

Kennedy, A. (1984) *The Psychology of Reading.* London: Methuen, pp. 160, £11·95 (h/b), £4·95 (p/b).

Lansdown, R. (1984). *Child Development.* London: Heinemann, pp. 256, £3·95.

Lawrence, J., Steed, D., and Young, P. (1984). *Disruptive Children: Disruptive Schools?* Beckenham: Croom Helm, pp. 263, £14·95.

Leff, H. L. (1984). *Playful Perception.* Vermont: Waterfront Books, pp. 161, $9·95 (softbound), $15·95 (hardbound).

Levy, P., and Gostein, H. (1984). *Tests in Education: A Book of Critical Reviews.* London: Academic Press, pp. 718, £20·00.

Maltby, F. (1984). *Gifted Children and Teachers in the Primary School 5-12.* London: Taylor and Francis, pp. 247, £13·95 (h/b), £8·95 (p/b).

Meddis, R. (1984). *Statistics Using Ranks.* Oxford: Basil Blackwell, pp. 449, £22·50.

Midgley, M. (1984). *Wickedness: A Philosophical Essay.* London: Routledge and Kegan Paul, pp. 224, £14·95.

Montgomery, D. (1984). *Evaluation and Enhancement of Teaching Performance.* Kingston Polytechnic, pp. 107.

Muller, D. J. (1984). *Remediating Children's Language.* Beckenham: Croom Helm, pp. 256, £14·95.

Muller, D. J., Blackman, D. E., and Chapman, A. J. (1984). *Psychology and Law.* Bognor Regis: John Wiley and Sons, pp. 472, £29·95.

Mussen, P. H., Conger, J. J., Kagan, J., and Huston, A. C. (1984). *Child Development and Personality.* London: Harper and Row, pp. 590, £9·95.

Ney, P. G., and Mulvihill, D. (1984). *Child Psychiatric Treatment.* Beckenham: Croom Helm, pp. 367, £35·00 (h/b), £12·95 (p/b).

Nisbet, J., and Shucksmith, J. (1984). *The Seventh Sense.* Edinburgh: SCER, pp. 17, £1·60.

Palmer, R., and Pope, C. (1984). *Brain Train: Studying for Success.* London, E. and F. N. Spoon, pp. 197, £4·95 (p/b), £12·95 (h/b).

Peins, M. (Ed.) (1984). *Contemporary Approaches in Stuttering Therapy.* Beckenham: Quest, pp. 304, $21·00.

Perera, K. (1984). *Children's Writing and Reading.* Oxford: Basil Blackwell, pp. 354, £8·95.

Rousseau, J. J. (edited by Allan Bloom) (1984). *Emile or On Education.* London: Harper and Row, pp. 501, £10·50.

Schreiber, F. R. (1984). *The Shoemaker: Anatomy of a Psychotic.* Harmondsworth: Penguin, pp. 432, £2·50.

Scraton, P., and Gordon, P. (1984). *Causes for Concern.* London: Pelican Original, pp. 287, £3·95.

Skynner, R., and Cleese, J. (1984). *Families and How to Survive.* London: Methuen, pp. 304, £3·50.

Stirling, E G. (1984). *Spelling Check-List: A Dictionary for Dyslexics.* Llandudno: St. David's College, £1·00.

Stones, E. (1984). *Psychology of Education.* London: Methuen, pp. 490, £5·95.

Thomson, M. (1984). *Developmental Dyslexia.* London: Edward Arnold, pp. 277, £12·50.

Tiberghien, G. (1984). *Initiation a la Psychophysique.* Paris: Presses Universitaires de France, pp. 152, 99F.

Tizard, B., and Hughes, M. (1984). *Young Children Learning.* London: Fontana, pp. 286, £2·95.

Tucker, I., and Nolan, M. (1984). *Educational Audiology.* Beckenham: Croom Helm, pp. 406, £25·00 (h/b), £13·95 (p/b).

Varma, V. P. (1984). *Anxiety in Children.* Beckenham: Croom Helm, pp. 222, £14·95.

Wheldall, K , and Merrett, F. (1984). *Positive Teaching: The Behavioural Approach.* Hemel Hempstead: Allen and Unwin, pp. 180, £12·00 (h/b), £5·50 (p/b).

Williams, P. (Ed.) (1984). *Special Education in Minority Communities.* Milton Keynes: Open University Press, pp. 138, £15·00 (h/b), £5·95 (p/b).

Child development and social policy: The life and work of Jack Tizard

Edited by **A. D. B. Clarke** *and* **B. Tizard**

This book is much more than a biography or posthumous tribute to a notable figure in British psychology. It has wide implications for behavioural scientists interested in the development of social policy in the light of empirical findings, providing a blue-print for applied social research.

Seventeen of Tizard's most notable contributions are reprinted here, together with two biographical chapters. The Editors have provided linking commentaries which place each contribution in context, and show how his research strategies developed. Papers are collected under four main headings: I. The mentally handicapped; II. Environmental influences; III. Services; IV. Epilogue.

In many ways ahead of his time, Tizard's researches repay close study. Their influence has been — and continues to be — considerable.

Pp 272 Price £7.95 casebound only ISBN 0 901715 18 2 Publication date April 1983

cheques with order to

The British Psychological Society

The Distribution Centre, Blackhorse Road, Letchworth, Herts SG6 1HN, UK

Call for papers

The British Journal of Developmental Psychology

Special issue on 'Language and Cognition in Early Social Interaction'

Papers are invited for this special issue of the journal to be published as part of Volume 4, 1986. Papers reporting original, empirical research will be especially welcome, and they should address one or both of the following topics: *(a)* The *acquisition* of either linguistic or cognitive skills within the social-interactive context; *(b)* the *deployment* of either linguistic or cognitive skills in the social-interactive context. All papers submitted for the special issue will be refereed in the usual way.

Closing date for submissions: 30 April 1985

Copies of submissions should be sent in triplicate to either of the two Guest Editors: Dr Martyn Barrett, Department of Psychology, Roehampton Institute, Digby Stuart College, Roehampton Lane, London SW15 5PH, *or* Dr Margaret Harris, Department of Psychology, Birkbeck College, University of London, Malet Street, London WC1E 7HX

The *British Journal of Developmental Psychology* is edited by Professor Peter Bryant. Volume 3 (1985), price £35.00 (US$ 64.50 overseas).

Orders to:

The British Psychological Society

The Distribution Centre, Blackhorse Road, Letchworth, Herts SG6 1HN, UK

Journals of The British Psychological Society

British Journal of Developmental Psychology
Edited by Peter Bryant, University of Oxford
Publishes empirical, conceptual and review articles relating to all areas of development *Launched in 1983*
Volume 3 (1985) (4 parts) US$64.50 (£35.00)

British Journal of Psychology
Edited by D E Blackman, University College, Cardiff
Reports on empirical studies, critical reviews of the literature and theoretical contributions
Volume 76 (1985) (4 parts) US$110.00 (£60.00)

British Journal of Medical Psychology
Edited by Sidney Crown, The London Hospital
Original contributions of knowledge in the area of those aspects of psychology applicable to medicine and related clinical disciplines including the psychotherapies are published
Volume 58 (1985) (4 parts) US$90.50 (£49.00)

British Journal of Social Psychology
Edited by A S R Manstead, University of Manchester
Publishes articles which contribute to the basic methodological issues confronting the discipline and includes theoretical and review papers, descriptions of applied social psychology in a variety of settings, research in growth areas of social psychology
Volume 24 (1985) (4 parts) US$64.50 (£38.00)

British Journal of Clinical Psychology
Edited by David A Shapiro, MRC/SSRC Social and Applied Psychology Unit, University of Sheffield
New findings, theoretical, methodological and review papers bearing on the whole field of clinical psychology are published Contributions combine scientific excellence with relevance to the concerns of practising clinicians
Volume 24 (1985) (4 parts) US$64.50 (£38.00)

British Journal of Mathematical and Statistical Psychology
Edited by Chris Leach, University of Newcastle upon Tyne
Publishes articles which have a clear reference to substantive psychological issues but with a greater mathematical or statistical or other formal aspect to their argument than is usually acceptable to other journals
Volume 38 (1985) (2 parts) US$79.50 (£41.00)

Journal of Occupational Psychology
Edited by Peter Warr, MRC/SSRC Social & Applied Psychology Unit, University of Sheffield
An international journal of industrial and organizational psychology which describes and interprets new research into people at work
Volume 58 (1985) (4 parts) US$79.50 (£41.00)
All seven journals ordered at the same time for one address US$500 (£273)

Invitation to readers
The British Psychological Society invites readers to become **Foreign Affiliates** of the Society The cost of Foreign Affiliateship is US$30 (£18)/yr Foreign Affiliates are entitled to receive the Society's Bulletin free (retail price US$25 (£15)) and to receive journals at member prices of US$10.00 (£6) per volume (for 1985) Foreign Affiliates are required to be members of a psychological society affiliated to IUPS

The British Psychological Society
The Distribution Centre, Blackhorse Road, Letchworth, Herts SG6 1HN, UK

British Journal of Developmental Psychology

Editor Peter Bryant, Watts Professor of Psychology, University of Oxford.

The immediate success of this new journal, launched in 1983, has demonstrated the central importance of its subject matter to the present-day study of psychology. The journal publishes full-length conceptual, empirical and review papers, as well as brief reports on work in progress, in the following areas. Development during childhood and adolescence, Early infant perceptual, cognitive and motor development; Abnormal development — the problems of handicaps, learning difficulties and childhood autism; Eductional implications of child development; Parent-child interaction; Acquisition of language; Social and moral development; Effects of ageing

Contributors to volumes 1 and 2 include such internationally recognized authorities as Gavin Bremner, A. D. B. Clarke, Antionette Krupski, Robert Hinde, Leon Mann, Susan Michie, David Olson, Michele Robert, Talib Rothengatter, James Russell, Michael Rutter, Sandra Scarr, Michael Schleifer, Anat Scher, Alan Slater, Peter Smith, Joan Tephly and Willam Yule.

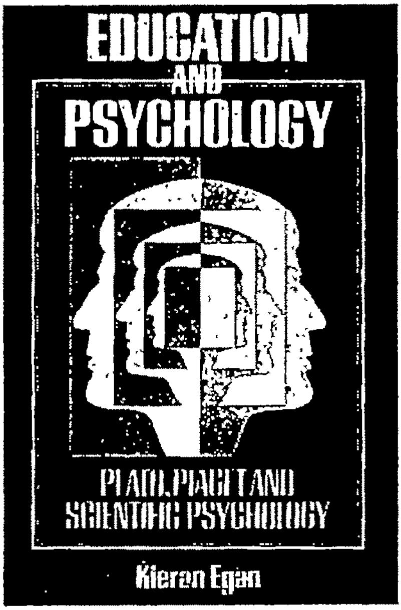
The journal has adopted a policy of devoting special issues to topics of particular interest within the field of developmental psychology and Part 4 1983 was a special issue devoted to the work of Neil O'Connor and colleagues at the MRC Developmental Psychology Unit, London. In addition, an issue of volume 3 1985 will be devoted to development in infancy, guest edited by George Butterworth and Paul Harris, and an issue of volume 4 1986 to language and cognition in early social interaction, guest edited by Martyn Barrett and Margaret Harris.

ISSN 0261-510X Volume 2 (1984) is still available for £33.00 (US$64.50); Volume 3 (1985) £35.00 (US$64.50)

Members and Foreign Affiliates of The British Psychological Society are entitled to purchase the journal for £6.00 (US$10.00) per volume.
Orders to

The British Psychological Society

The Distribution Centre, Blackhorse Road, Letchworth, Herts SG6 1HN, UK



Education and Psychology

Plato, Piaget and scientific psychology

KIERAN EGAN

'Psychology of education' has long held a place in the curriculum for training teachers; but what implications can psychological theory legitimately have for educational practice? In this clearly argued and lively book Kieran Egan makes a direct attack on the current role of psychology in education, showing important differences between psychologists' and educators' interests in topics such as learning, motivation and development, and questioning the validity of many of Piaget's fundamental ideas.

224 pages
Hardback 0 416 38100 6 £10.50
Paperback 0 416 38110 3 £4.95

METHUEN
11 New Fetter Lane, London EC4P 4EE

Centre for the Study of Management Learning

CSML

Fourth National Conference on Education and Development in Organisations

OPEN AND DISTANCE LEARNING IN EDUCATION AND DEVELOPMENT:

Issues in Theory and Practice

17th–20th September 1985
University of Lancaster

The Conference is for those with an interest in education and development in any type of organisation. The aim is to bring together people from different fields to explore the issues of theory and practice in open and distance learning applied to education and development.

Contributions invited of any type (presentations, papers, workshops, etc).

Further details from Christine Dytham, Conference Secretary, C.S.M.L., Gillow House, University of Lancaster, Lancaster LA1 4YX.
Tel: (0524) 65201 Ext. 4698/4854

PUBLICATIONS OF THE BRITISH PSYCHOLOGICAL SOCIETY

British Journal of Psychology
British Journal of Medical Psychology
Journal of Occupational Psychology
British Journal of Social Psychology
British Journal of Clinical Psychology
British Journal of Developmental Psychology
British Journal of Mathematical and Statistical Psychology
The British Journal of Educational Psychology

The *British Journal of Psychology*, the *British Journal of Medical Psychology*, *Journal of Occupational Psychology*, the *British Journal of Developmental Psychology*, the *British Journal of Social Psychology* and the *British Journal of Clinical Psychology* appear in quarterly parts. The 1985 subscription prices, payable in advance, are £60·00 (US$110·00) for the *British Journal of Psychology*; £38·00 (US$64·50) for the *British Journal of Social Psychology*; £38·00 (US$64·50) for the *British Journal of Clinical Psychology*; £35·00 (US$64·50) for the *British Journal of Developmental Psychology*; £49·00 (US$90·50) for the *British Journal of Medical Psychology*; and £41·00 (US$79·50) for the *Journal of Occupational Psychology*.

The *British Journal of Mathematical and Statistical Psychology* is issued twice a year, in May and November. The subscription rate, payable in advance, is £41·00 per volume post free (US$79·50 overseas).

Subscriptions to any of these journals may be sent to a bookseller or subscription agent, or to The British Psychological Society, The Distribution Centre, Blackhorse Road, Letchworth, Herts SG6 1HN, UK. A special price of £273·00 (US$500·00) for all seven journals going to the same address is offered for 1985.

The British Journal of Educational Psychology is issued three times a year, in February, June and November. The prepaid subscription is £18·00 a year, post free, or £6·50 each number. The prepaid subscription price in the USA is $40·00 post free. Subscriptions should be sent to the Scottish Academic Press, 33 Montgomery Street, Edinburgh EH7 5JX.

MANUSCRIPTS

For the *British Journal of Psychology* should be sent to Professor D. E. Blackman, Department of Psychology, University College Cardiff, PO Box 78, Cardiff CF1 1XL.

For the *British Journal of Medical Psychology* should be sent to Dr Sidney Crown, Department of Psychiatry, The London Hospital, Whitechapel, E1 1BB.

For the *Journal of Occupational Psychology* should be sent to Dr D. Guest, Department of Industrial Relations, London School of Economics and Political Science, Houghton Street, London WC2A 2AE.

For the *British Journal of Social Psychology* should be sent to Dr A. S. R. Manstead, Department of Psychology, The University, Manchester M13 9PL.

For the *British Journal of Clinical Psychology* should be sent to Dr D. Shapiro, MRC/SSRC Social and Applied Psychology Unit, Department of Psychology, The University, Sheffield S10 2TN.

For the *British Journal of Mathematical and Statistical Psychology* should be sent to Dr C. Leach, Department of Psychology, University of Newcastle upon Tyne, Newcastle upon Tyne NE1 7RU.

For the *British Journal of Developmental Psychology* should be sent to Professor P. E. Bryant, Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford OX1 3UD.

For *The British Journal of Educational Psychology* should be sent in duplicate to Professor Geoffrey Brown, School of Education, University of East Anglia, Norwich NR4 7TJ.

Contributors are advised to consult *Suggestions to Contributors* issued by the Standing Committee on Publications of the British Psychological Society and obtainable, price £2·00 post free, from The British Psychological Society, St. Andrew's House, 48 Princess Road East, Leicester LE1 7DR.

All manuscripts should be submitted in duplicate.

OFFPRINTS

Contributors on application receive fifty copies of their papers free. Additional copies may be obtained in units of fifty copies.

REPRINTS

The *Journal* is available in reprint from Swets and Zeitlinger B.V., P.O. Box 810, 2160 SZ Lisse, The Netherlands; Vols. 21-29, 31-47.

SUBSCRIPTIONS TO THE BRITISH PSYCHOLOGICAL SOCIETY

The subscription for Fellows, Associates and Graduate Members is £30·00.

They are entitled to purchase any or all of the Society's Journals at the special rates of £6·00 per Journal per annum.

Information on membership and activities may be obtained from the Secretary, The British Psychological Society, St. Andrew's House, 48 Princess Road East, Leicester LE1 7DR.

BRITISH JOURNAL OF EDUCATIONAL PSYCHOLOGY,

VOL. 55, PART 1, 1985

CONTENTS

ARTICLES:

ZAKAY, D.	The Influence of Information and Daily Contact on Children's Attitudes Towards Aphasic Children	1
LITTLE, A. W.	The Child's Understanding of the Causes of Academic Success and Failure: A Case Study of British Schoolchildren	11
KLINE, P., and COOPER, C.	Rigid Personality and Rigid Thinking	24
YATES, G. C. R., HANNELL, G., and LIPPETT, R. M.	Cognitive Slippage, Test Anxiety, and Responses in a Group Testing Situation	28
JOHNSON, M. R. D., BEWLEY, B. R., BANKS, M. H., BLAND, J. M., and CLYDE, D. V.	Schools and Smoking: School Features and Variations in Cigarette Smoking by Children and Teachers	34
HARDING, L. M., BEECH, J. R., and SNEDDON, W.	The Changing Pattern of Reading Errors and Reading Style from 5 to 11 Years of Age	45
FREEBODY, P., and TIRRE, W. C.	Achievement Outcomes of Two Reading Programmes: An Instance of Aptitude-Treatment Interaction	53
KYRIACOU, C. and PRATT, J.	Teacher Stress and Psychoneurotic Symptoms	61
WHELDALL, K., MERRETT, F., and BORG, M.	The Behavioural Approach to Teaching Package (BAT-PACK): An Experimental Evaluation	65

RESEARCH NOTES:

GIBBS, J., and NORWICH, B.	The Validity of a Short Form of the Coopersmith Self-Esteem Inventory	76
LOW GUAT TIN and TEASDALE, G. R.	An Observational Study of the Social Adjustment of Spina Bifida Children in Integrated Settings	81
BEVERIDGE, M.	The Development of Young Children's Understanding of the Process of Evaporation	84
YOUNGMAN, M. B., and SZADAY, C. S. B.	Further Validation of the Behaviour in School Inventory	91
CONFERENCE ABSTRACTS		49
PUBLICATIONS RECEIVED		108

26 MAY '86

CALCUTTA UNIVERSITY

UK ISSN 0007-0998

THE BRITISH JOURNAL OF EDUCATIONAL PSYCHOLOGY

370.1505
B777

370.1505
B777

JUNE 1985

VOL. 55, PART 2

CALCUTTA

THE
BRITISH JOURNAL
OF
EDUCATIONAL PSYCHOLOGY

Edited by
Professor GEOFFREY BROWN

Assistant Editors:
Dr. R. L. BURDEN
Dr. C. W. DESFORGES
Dr. D. FONTANA

Sub-Editor:
Dr. S. CUNNINGHAM

Book Review Editor:
Dr. C. W. DESFORGES

Advised by the Editorial Committee:
Dr. H. BELOFF
Professor M. B. SUTHERLAND
Professor J. WRIGLEY

Vol. 55, Part 2

The British Journal of Educational Psychology is issued by the British Psychological Society once a term, in February, June and November The subscription is £18 00 a year post free, payable in advance, or £6 50 each number.

Subscriptions and Orders should be sent to Scottish Academic Press, 33 Montgomery Street, Edinburgh, EH7 5JX Members of the British Psychological Society receive the *Journal* on special terms

Papers for publication in duplicate should be sent to Professor Geoffrey Brown, School of Education, University of East Anglia, Norwich NR4 7TJ Guidance to contributors on the preparation of papers for the *Journal* is given at the end of the November number each year.

All *Books for Review* for *The British Journal of Educational Psychology* should be sent to. Dr. C W. Desforges, School of Education, University of East Anglia, Norwich NR4 7TJ

Advertisements for this *Journal* should be sent to Scottish Academic Press

Published by
SCOTTISH ACADEMIC PRESS, 33 MONTGOMERY STREET, EDINBURGH, EH7 5JX,
for the British Journal of Educational Psychology, Ltd.

All rights reserved

© 1985 The British Journal of Educational Psychology

PRINTED IN GREAT BRITAIN BY CLARK CONSTABLE, EDINBURGH, LONDON, MELBOURNE

urnal o

Br. J. educ. Psychol., 55, 111-118, 1985

370.156 S
B 777

SOCIAL BEHAVIOUR OF HEARING-IMPAIRED AND NORMALLY-HEARING PRESCHOOLERS

BY RACHEL LEVY-SHIFF AND MICHAEL A. HOFFMAN
(Bar-Ilan University, Ramat-Gan, Israel)

SUMMARY The relationship between social behaviour and hearing loss was studied in Israeli preschool-age children Free play interactions were observed in mainstreamed settings among 36 children with varying degrees of hearing loss (severe loss, profound loss) and normal hearing Social competence and hearing impairment were inversely related: the profoundly impaired were markedly less skilful than the normal hearing, while the severely impaired differed only moderately. However, the hearing impaired as a group displayed greater self-segregation tendencies to interact with those of similar hearing status than was displayed by the normal hearing These results are discussed in light of extant research and theory on mainstreaming.

INTRODUCTION

VERBAL communication has a major role in the initiation and maintenance of social interactions, as well as in the organisation and structure of the social situation (Piaget, 1926; Mueller, 1972; Garvey and Hogan, 1973). As such, hearing-impaired children may be expected to manifest primary development difficulties not only in language-related social skills but also in overall social competence. Indeed, among older hearing-impaired children, social developmental lag and malfunctioning have been found in several studies (for review, see Meadow, 1975). Inadequacies have included inaccuracy of self-perceptions (Craig, 1965), misinterpretation of social cues (Odom *et al.*, 1973), communication dysfunction (Heider and Heider, 1971; Hoemann, 1972; Schlesinger and Meadow, 1972), and social immaturity (Myklebust, 1960).

Research on the effect of functional disability on social behaviour and development among preschool-age children is less clear than that for their older peers. Surprisingly, very few studies have investigated the social behaviour of young hearing-impaired children, and these have found little evidence for a development lag in competence. Vandell and George (1981), in studying the social initiations of hearing-impaired (HI) and normally-hearing (NH) preschoolers in a laboratory setting, concluded that the communication skills of the two groups were basically equivalent Similar conclusions of interactional parity were drawn by Arnold and Thremblay (1978), based on observations of HI and NH preschoolers' social behaviour in a natural free play setting. Except for minor differences in vocalisation frequencies, the social behaviour of HI preschoolers was not significantly different from that of NH peers.

Close inspection of these studies suggests that the social behaviour of HI preschoolers may differ more than these researchers suggest. Methodological limitations appeared to have hampered both studies. Vandell and George (1981), for example, focused exclusively on initiation without regard to the wider variety of behaviours involved in social interaction. Moreover, their study was conducted in an artificial setting between arbitrarily contrived dyads in which interaction was almost a certainty. Finally, they did not discriminate between those with either profound or moderate hearing loss, a distinction with important ramifications for expected communication impairment. While the work by Arnold and Thremblay successfully circumvented these problems in behavioural focus, setting, and hearing loss distinction, their statistical power to discern differences was greatly reduced by small sample size (N = 12). These limitations suggest the need for additional research concerning the impact of hearing impairment on preschool-age children's social skills.

One additional issue arising from research on older hearing-impaired is their exposure to segregational reactions (Elser, 1959; Justman and Moskowitz, 1975; Plaut and Kravetz, 1982). With regard to younger children, Vandell and her associates (Vandell and George, 1981; Vandell *et al.*, 1982), found that HI were likely to have their initiations rejected by NH or receive ineffectual initiations from them. Arnold and Thremblay (1978) also found segregational tendencies toward the HI among NH preschoolers. The NH tendency to avoid the HI peers may lead to deficiencies in the quality of socialisation experiences. However, little work has attempted to assess the degree to which these segregational tendencies are associated with the relative hearing loss and social skills of the HI involved.

The present study was designed to investigate the impact of differing levels of hearing impairment on social competence and segregation among preschool-age children in a free play setting. It was expected that children suffering higher levels of hearing loss would show reduced social competence and integration. This expectation was based on the view that decrements in the ability to take part in verbal communication should lead to reduced capacity to organise and maintain social interaction with normal-hearing children.

METHOD

Subjects

The sample consisted of 12 profoundly hearing-impaired (PHI), 12 severely hearing-impaired (SHI), and 12 normal-hearing (NH) preschoolers, with equal numbers of boys and girls in each group. Severe impairment referred to a loss of 70 dB-90 dB in the better ear for frequencies important for hearing speech (500-2000 cps), while profound impairment referred to a loss of 90 dB or more. Impairment was either congenital or had occurred in early infancy before the stage of language acquisition. However, no child suffered systemic neurological impairment. Moreover, requirements for acceptance to the kindergarten programme included that of normal intellectual performance.

Age ranges and means were similar for the NH (49 to 67 months, M = 60 months), SHI (49 to 65 months, M = 58 months), PHI (48 to 64 months, M = 59 months). Children came from predominantly middle-class families of Ashkenazic Jewish origin; chi-square tests of ethnic origin and SES background variables did not reveal significant between-group differences.

The HI and NH children were classmates in five integrated kindergartens Two or three children were randomly selected for the three groups from the class lists of those who met required criteria study in each kindergarten (age, normal development, intact family, SES). The children had been in school together for five months at least prior to the beginning of the study.

The PHI and SHI children were participants in the second stage of the municipal Tel-Aviv integration programme for hearing-impaired children. In the first stage, young hearing-impaired children between the ages of about 2·6 and 4·0 years had attended regular neighbourhood nurseries on an individual basis four days a week. For another two days a week they had attended the special Micha (Centre for Deaf and Mute Children) nursery for hearing-impaired children where they were given early auditory-oral training individually and in groups, as well as a variety of group tutoring activities. The means of communication were exclusively oral, without reliance on manual (sign language) or total communication techniques. Intervention consisted of early amplification, training in verbal language and upgrading of residual listening skills when present. In the second stage, starting at 4·0 to 4·6 years of age, the hearing-impaired children were placed in integrated kindergarten classes, located in regular kindergartens rather than special schools or centres for the hearing-impaired. Each class had eight hearing-impaired and 25 normal-hearing children under the tutelage of two teachers, at least one of whom was a trained teacher of hearing-impaired or exceptional children. All children attended the regular Israeli regimen of six days a week, five hours a day. Curriculum, based on an open-classroom and activity-centre format, was the same as that in all Israeli kindergartens. The HI children participate jointly with the NH children in all activities — except the group activity focused on teaching school-room subjects by exclusively verbal means (about 30 to 60 minutes a day). This curriculum calls for an additional teacher to give the HI children individual language lessons twice a week. The means of communication remain exclusively oral. This programme is designed to make possible the maximal integration of HI children in a normal hearing environment. The special features of this programme — early start, exclusively oral training, early exposure to NH peers, and the absence of highly salient gesturing aids to communication — are intended to reduce to a minimum the interference of secondary factors that may hinder the hearing-impaired child's socialisation and the development of his social skills.

Procedure and measures

Hearing-impaired children and normal-hearing classmates were observed by two graduate students of the Special Programme for Communication Disorders in Tel-Aviv University. All observations were made during normal free play periods in kindergarten play areas with a similar assortment of preschool play materials. The only instructions given to teachers were to refrain from contact with the child under observation. During observation, the observer tracked the behaviour of the target child at close range and refrained from personal interaction. Subject-observer interference was further reduced by observers' classroom visits prior to formal observation as well as by initial familiarisation with the subject at the beginning of each session.

Four 15-minute observation periods were conducted over a period of one month for each target child. Behaviour was time sampled on a checklist of 19 behaviours, with one tally being made for the first precoded behaviour occurring during each 15-second observation time unit. Thus, a total of 240 behavioural time units were collected for each child.

During periods of interaction, the observers also recorded the hearing status of the target child's partner. However, due to the number of children involved, observers were not always aware of the relative degree of hearing impairment. Therefore, observers only noted whether the partner was hearing-impaired or not, without distinguishing between profound and severe levels.

Each target child was observed simultaneously by a primary and secondary observer, with the pair of observers alternating roles between sessions. The primary's tallies served as the data base for analysis, while those of the secondary observer were used to assess reliability. Pearson product moment correlations between the two observers' scorings appear after each observed behaviour in the list given below.

The 19 behaviours selected for coding were culled from research in social competence with peers (Lieberman, 1977) and dealt with interactional ability, affect, and non-interaction. Behaviours reflecting *ability and willingness to interact with playmate* included: initiation of interaction that elicited positive response, 0·93; positive response to playmate's initiation, 0·89; verbal request positively received by partner, 0·83; conversational sequences, 0·94; giving, 0·95; communicative gestures, 0·91; sharing, 0·85; taking toy from a playmate without encountering resistance, 0·94. *Expressions of positive or negative affect* included: smile or laughter, 0·96; crying, 0·98; physical aggression, 0·95; violent treatment of toy, 0·89; verbal aggression and threats toward playmate, 0·92; struggle to take toy from a playmate, 0·91. *Absence of social* *interaction* included: staring at playmate without participating in his/her activity (non-participation attending), 0·93; solitary play, 0·96; wandering around, 0 97; self-directed activity (e.g , automanipulation), 0·90; other non-interactive behaviour, 0·85.

A composite measure reflecting the total amount of time in social interaction was computed from the sum of all behaviour unit tallies other than the five reflecting absence of social interaction.

Data analysis

Children's scores within each of the three groups were compared across the five kindergartens. In the absence of significant differences, data were pooled across kindergartens in subsequent analyses

The study comprised a 3 (hearing status of target child) × 2 (hearing status of partner) design with partner's hearing status treated as a within factor. As such, 3 × 2 ANOVAs of behavioural frequencies seemed a preferential method of analysis. Yet the ANOVA rests on the assumption that target children have equal opportunity to interact with both HI and NH partners when assessing both the main effect and interaction effect involving partners' hearing status (Otherwise, lower rates of interaction may simply be due to fewer opportunities for contact.) However, children actually had differing numbers of possible NH or HI partners (25 *vs*. 8 respectively).

In order to control for this issue, chi-square analyses (Siegel, 1956) were used to supplement the ANOVA techniques. In these, the ratio of observed interactional frequencies for partners was compared to that predicted from their relative numbers in the play setting. To assess the main effect of partners, this chi-square comparison was conducted first among target children as a whole. Separate chi-square analyses for targets were used to assess possible interactions of target and partner. (The possible use of correction factors involving average number of behaviours per target person was eliminated due to a lack of data regarding the number of targets involved in each interaction.)

In sum, all behavioural measures involving social interaction were subjected to 3 × 2 ANOVAs as well as by chi-square tests. Since non-interactive measures did not include partners, these were analysed by simple one-way ANOVAs for target child effects. In all cases, ANOVAs were followed by Scheffé tests for *post hoc* comparisons of means

RESULTS

Behavioural time units for the three groups of target children are broken down by status of hearing partner in Table 1.

The two-way ANOVA of total time units of social interaction revealed significant differences between the target children, $F(2,33)=3{\cdot}56$, $P<0{\cdot}04$. Subsequent Scheffé tests indicated that the PHI were engaged significantly less in social interaction (118·3 time units, 49·3 per cent of total observation time) than either the SHI (158 1 time units, 65·8 per cent of total observation time) or the NH (163·5 time units, 68·1 per cent of total observation time). Separate ANOVAs for each behavioural category revealed that the interactional differences between target children fell in the areas of initiation, $F(2,33)=4{\cdot}43$, $P<0{\cdot}02$, and conversation, $F(2,33)=5{\cdot}16$, $P<0{\cdot}01$. The Scheffé tests indicated that the PHI and the SHI initiated interactions that elicited a positive response from peers less than the NH children. Moreover, the PHI were found to be engaged in conversation less than the SHI children, and the latter less than the NH children. No significant differences appeared between target children in the remaining categories of social ability or affect.

No significant main effect was found for partner in either the ANOVA of the total

TABLE 1

MEAN SOCIAL BEHAVIOUR TIME UNITS OF SEVERELY HEARING IMPAIRED, PROFOUNDLY HEARING IMPAIRED, AND NORMAL HEARING PRESCHOOLERS TOWARD NORMAL AND HEARING IMPAIRED PARTNERS

Target (T)	PHI		SHI		NH		F[1]		
Partner (P)	NH	HI	NH	HI	NH	HI	T	P	T×P
Social willingness and ability									
Initiation	3·83	15·10	5·30	19·12	32·91	0·63	4·43*	—	26·10***
Positive response	7·05	18·31	6·02	21·31	26·75	2·95	—	—	17·98***
Verbal request	1·18	4·24	5·23	7·43	8·26	1·33	—	—	9·97***
Conversation	7·05	20·67	8·07	35·12	52·93	5·86	5·16**	—	18·68***
Giving	1·61	1·72	2·13	3·87	3·13	0·75	—	—	9·98***
Gestures	0·61	8·81	3·08	9·13	6·06	0·72	—	3·91*	14·32***
Sharing	0·05	2·59	1·32	2·41	2·86	0·53	—	—	9·45***
Taking a toy	1·22	5·24	1·73	4·13	4·46	0·00	—	—	18·22***
Affect									
Smile, laughter	1·13	3·25	0·43	2·25	4·33	0·00	—	—	5·04**
Crying	0·00	1·27	1·28	1·56	0·40	0·00	—	—	2·87'
Physical aggression	0·25	3·00	2·30	3·52	3·60	0·00	—	—	8·31***
Violence of toy	1·42	2·67	0·00	3·97	0·93	0·00	—	5·38*	5·21**
Verbal aggression	0·00	1·52	0·69	2·89	2·83	0·06	—	—	4·04*
Struggle over toy	1·21	3·35	1·32	2·51	1·26	0·00	—	—	3·28*
Non-social behaviour									
Staring	10·61	31·44	7·51	14·28	21·73	0·02	3·59*	—	10·42***
Solitary play[2]	56·38		40·31		48·33		—		
Wandering[2]	11·22		12·63		3·53		3·75*		
Self-directed[2]	5·47		3·25		3·20		—		
Others[2]	2·54		4·33		0·93		—		

NOTE PHI = Profoundly hearing impaired; SHI = Severely hearing impaired, NH = Normal hearing; HI = Hearing impaired (profound & severe)

[1] Significant Fs in 3 × 2 ANOVA (N = 36)

[2] As a non-interactive behaviour, no breakdown is offered for partner and F reflects a one-way ANOVA for target

' $P<0·10$, * $P<0·05$; ** $P<0·01$, *** $P<0·001$

interaction time or in the subsequent χ^2. However, both the ANOVAs and χ^2 analyses for the separate categories revealed significant differences in regard to gestures, $F(1,33) = 3·91$, $P < 0·05$, and violent toy use, $F(1,33) = 5·38$, $P < 0·02$. More gestures and more frequent violent toy use appeared in exchanges with HI rather than NH partners.

A significant interaction effect between target and partner appeared in both the ANOVA and χ^2 of total interaction time, $F(2,33) = 32·83$, $P < 0·001$. Inspection of Table 1 suggests that children interacted more with partners whose hearing status matched their own. In subsequent ANOVAs, significant interactive effects were also found for each of the 14 specific behavioural categories related to social ability and affect. The parallel χ^2 analyses indicated that both impaired and normal-hearing children preferred to initiate, converse, and positively respond to like-hearing partners. However, only the PHI and SHI, but not the NH, displayed partner preference in regard to the 11 remaining behavioural categories.

As noted above, PHI entered into interaction less than the two other groups. Analysis of the separate categories of non-social behaviour revealed significant differences in target children's activities during periods not involving direct interaction. A 3 × 2 ANOVA of staring indicated both a significant main effect for target children, $F(2,33) = 3·59$, $P < 0·04$, and an interactive effect between target partner, $F(2,33) = 10·42$, $P < 0·001$. χ^2 tests revealed that all children preferred to stare at children of similar hearing status, yet PHI were found in Scheffé tests to stare more than either SHI or NH children.

In the one-way ANOVAs of the three especially non-interactive categories, significant differences appeared in wandering, but not in either solitary play or self-directed activity. Both profoundly and severely HI wandered significantly more than the NH children, $F(1,33) = 3{\cdot}75$, $P < 0{\cdot}05$.

In order to assess qualitative differences in the profile of skills employed, the 3 × 2 ANOVAs were performed on the time spent in different forms of interaction as percentages of total time in social interchange. No gross differences were revealed in the relative use of specific skills between the impaired and the normal-hearing. The single exception to this generalisation was that HI spent significantly more of their social interaction time in communicative gesturing than NH, $F(2,33) = 4{\cdot}85$, $P < 0{\cdot}03$.

DISCUSSION

Review of the findings suggests that PHI are markedly less socially competent than NH children, while SHI children are only moderately so. The PHI spent significantly less time in overall contact with others, with a concordant drop-off in almost all specific social skill areas. In contrast, the SHI children had almost as much social contact as the NH. The single differences between the SHI and normal children involved the former's greater reliance on gestures and the latter's involvement in verbal conversation.

Analysis of the PHI children's activities when not in contact with others suggests that these children have lower social competence, but not lower social interest. Nearly 55 per cent of the time these children were not actively involved with peers was invested in passively watching them. Staring seems to borderline between interaction and withdrawal, reflecting a passive involvement. Thus, PHI children's staring would seem to indicate that they have social interests equalling those of NH children, but lack the specific skills for successfully initiating and maintaining contact.

In contrast to the PHI, SHI displayed a gross social competence basically equivalent to the NH However, specific skill differences of note did arise. SHI, as well as the PHI children, relied more heavily on gestures, while making less use of verbal tools than NH children. Although these differences may appear at first glance to be stylistic, issues arise as to their meaning in terms of possible social communication deficiency.

Non-verbal communication is often thought to be less powerful and efficient than the verbal mode (Meadow, 1975). Indeed, the increased appearance of violent interchanges among both the PHI and SHI would seem to reflect instances of communication failure and frustration in peer interaction. Thus, the rough equivalence in social competence between SHI and NH does not deny the presence of potential functional speech problems associated with hearing impairment.

Overall, the present findings among HI extend those of previous research and amplify the importance of distinctions in hearing levels. The relationship between social competence and hearing loss found in this study's preschool-age population parallels that found in work with older school-age children (Reich *et al.*, 1977). Methodologically, this finding suggests that previous conclusions on social parity between normal and hearing-impaired preschoolers may have been due, in part, to the level of hearing loss surveyed. Indeed, review of Vandell and George's (1981) work indicates that a number of their sample fell into the severe rather than profound groups of impairment. Theoretically, it suggests that social functioning and social skills development are related to the degree of hearing loss. Pragmatically, it underlines the need for greater attention to the issue of relative hearing loss in programming for the hearing impaired (Berg, 1976).

A second focus of study results was the widespread appearance of segregational tendencies among the impaired as well as the normal hearing children. Both groups of children interacted predominantly with partners of similar hearing status. This segregational tendency appeared in almost every observational measure of interaction among the HI, and arose significantly for the NH in regard to the high frequency behaviours of conversation and initiating contact with others. The segregational tendencies of the hearing impaired became notable once the differing proportions of normal and hearing-impaired children in the class were controlled for.

The segregational tendencies apparent in this sample partially coincide with those found among PHI preschoolers by Arnold and Thremblay (1978). They reported strong segregational tendencies among NH children, but only slight and not significant tendencies among the PHI. Differences between the two studies may derive from the fact that preschoolers in the present study had significant experience with each other for at least two years previous to observation. Thus, their segregational tendencies may have simply reflected previous friendship and preferential bonds. Moreover, the hearing-impaired children in our study may have been more conscious of their handicap, due to the intensive mainstreaming programme than were the children in Arnold and Thremblay's study who participated in a low key programme (Higgins, 1982). These explanations, however, do not explain the reasons underlying normal hearing children's segregational tendencies, both in this study and in Arnold and Thremblay's (1978) work.

It is important to note that the present study's findings offer little support for the naïve belief that the segregation of HI children arises from their purported social ineptness. If segregation were a function of mere social competence, one would have expected SHI to segregate less than the more inept PHI. However, no differences were found in segregational tendencies between these two groups.

Reference to the wider research literature (e.g., Gellman, 1959; Bateman, 1964; English, 1972; Katz *et al.*, 1973; Plaut and Kravetz, 1982) suggests that segregation of the HI may be due to their being viewed as different rather than inept. Segregational patterns have appeared among many groups of children, not all of whom lacked social skills (eg., the physically handicapped (Force, 1956) and culturally deprived children (Feitelson *et al.*, 1972)). As such, hearing impairment would seem to be only one of many factors upon which children make ingroup/outgroup distinctions underlying friendship. As Asher (1978) noted, other "difference" factors influencing reciprocal relations include sex, race, name, appearance, disability, attitudes and beliefs.

Before addressing the implications of this study, it is necessary to note that interpretation of these results requires caution. The study sample was drawn from a unique programme and thus the trends observed may be outcomes of the specific programme features. Moreover, it is not yet clear how much the patterns observed in the play periods characterised these children's behaviour in class or outside of school. These limitations on generalisability require future replication and extension on the present work.

Yet, if replicated, the results of this study have major implications for the formation of mainstreaming programmes for the HI. First, the inverse relationship between social competence and hearing loss suggests that the PHI population does require general social training. However, the SHI would seem to require more focused programmes dealing with only a limited range of social competencies. Second, the segregational tendencies reflected in this sample suggest that mainstreaming efforts must go far beyond simple integration of the child into the classroom. Even though children in the present study were involved in a highly intensive integration effort, the signs of segregation were clear and striking. Further research is needed to identify the source of these segregational tendencies in order to design effective programmes for their amelioration.

REFERENCES

ARNOLD, D., and THREMBLAY, A. (1978). Interaction of deaf and hearing preschool children. *J. Communic. Disorders*, 12, 245-251.

ASHER, S. R. (1978) Children's peer relations In LAMB, M E. (Ed.), *Social and Personality Development* New York. Holt, Rinehart and Winston.

BATEMAN, B. (1964). Sighted children's perceptions of blind children's abilities. *Except. Children*, 29, 42-46.

BERG, F. S (1976) *Educational Audiology: Hearing and Speech Management* New York Grune and Stratton

CRAIG, H. B. (1965) A sociometric investigation of self-concept of the deaf child *Am Annals of the Deaf*, 110, 456-478.

ELSER, R (1959). The social position of hearing handicapped children in regular grades. *Except. Children*, 25, 305-309.

ENGLISH, R W. (1972). *Correlates of Stigma towards Physically Disabled Persons* New York: Syracuse University Press

FEITELSON, D., WEINTRAUB, S , and MICHAELI, O. (1972). Social interactions in heterogeneous preschools in Israel *Child Dev* , 43, 1249-1259.

FORCE, D (1956). Social status of physically handicapped children. *Except. Children*, 23, 104-107.

GARVEY, C., and HOGAN, R (1973). Social speech and social interactions: egocentrism revisited. *Child Dev.*, 44, 562-568.

GELLMAN, W. (1959) Roots of prejudice against the handicapped *J Rehabilitation*, 25, 4-16.

HEIDER, F , and HEIDER, G. M (1971) Studies in psychology of the deaf *Psychol Monogr.*, 53, (no. 242).

HIGGINS, P C (1982) *Outsiders in a Hearing World* Beverly Hills, Calif . Sage Publications

HOEMANN, H. W (1972). The development of communication skills in deaf and hearing children *Child Dev.*, 43, 990-1003.

JUSTMAN, J , and MOSKOWITZ, L. (1975) *The Integration of Deaf Children in Hearing Classes* (Publication no 4) New York: Board of Education, Bureau of Educational Research.

KATZ, I , GLASS, P. C , and COHEN, S. (1973) Ambivalence, guilt, and scapegoating of minority group victims *J. exp. soc. Psychol.*, 4, 423-436.

LIEBERMAN, A F. (1977) Preschoolers' competence with a peer relations with attachment and peer experience *Child Dev.*, 48, 1277-1287

MEADOW, K P (1975). The development of deaf children In HETHERINGTON, M. E. (Ed.), *Review of Child Development Research* (Vol. 5) Chicago: University of Chicago Press.

MUELLER, E. (1972) The maintenance of verbal exchanges between young children *Child Dev.*, 43, 930-938.

MYKLEBUST, H R. (1960) *The Psychology of Deafness. Sensory Deprivation, Learning, and Adjustment* New York. Grune and Stratton

ODOM, P. P , BLANTON, R. L., and LAUKHUF, C (1973). Facial expressions and interpretation of emotion-arousing situations in deaf and hearing children *J. abnorm. Child Psychol*, 1, 139-151.

PIAGET, J (1926). *The Language and Thought of the Child* New York. Harcourt and Brace.

PLAUT, A., and KRAVETZ, S. (1982). Influence of kind of contact and teachers' attitude on social acceptance of deaf students by their hearing classmates. *Megamot*, 24, 294-302 (Hebrew)

REICH, C., HAMBLETON, D., and HOULDIN, B. K. (1977). The integration of hearing impaired children in a regular classroom *Am Annals of the Deaf*, 122, 534-543.

SCHLESINGER, H. S , and MEADOW, K. P (1972) *Sound and Sign Children's Deafness and Mental Health.* Berkeley University of California Press.

SIEGEL, S (1956). *Nonparametric Statistics for the Behavioral Sciences.* New York: McGraw-Hill.

VANDELL, D. L., ANDERSON, L. D., EHRHARDT, G., and WILSON, K. S. (1982) Integrating hearing and deaf preschoolers an attempt to enhance hearing children's interaction with deaf peers. *Child Dev* , 53, 1354-1363

VANDELL, D. L , and GEORGE, L B. (1981). Social interactions in hearing and deaf preschoolers: success and failures in initiations. *Child Dev.*, 52, 627-635

(*Manuscript received* 12th December, 1984)

Br. J. educ Psychol., 55, 119-123, 1985

TAKING CONTROL OF STRESS IN TEACHING

BY D. A. WOODHOUSE*, E. HALL AND A. D. WOOSTER
(*School of Education, University of Nottingham*)

SUMMARY An evaluation of two 6-day in-service courses in human relations for teachers, using a structured diary before, during and after the course, and a follow-up interview, produced data on teacher stress which was qualitatively different from earlier research using questionnaires. The courses were experiential and were perceived as stressful. In the follow-up diaries and interviews there was evidence that training in human relations produced changes in behaviour and a reduction in stress.

INTRODUCTION

A NUMBER of recent studies have shown that teachers in English schools experience their work as stressful in varying degrees. Dunham (1976) producess evidence that the experience of stress increased with the reorganisation of secondary education in the U.K., and later (1980) found that both English and German teachers perceived 'poor staff communications' and 'disruptive pupil behaviour' as major stress situations. Kyriacou and Sutcliffe (1978) identify four main sources of stress which they describe as 'pupil misbehaviour', 'poor working conditions', 'time pressures' and 'poor school ethos'. Fletcher and Payne (1982) also found reports of high levels of stress among teachers who identified 'demands made by individual children', 'not having enough time' and 'the continual need to be creative' as the main job demands.

The data supporting these findings are based on questionnaire responses, with the exception of Dunham (1980), who added a half hour interview. Smilansky (1984) related self-reported levels of stress in Israeli teachers to ratings by principals, other teachers, parents and pupils regarding their quality as teachers. The 'better' teachers reported higher levels of stress. Smilansky suggests that the 'better' teacher has the professional confidence to admit to the experience of stress, which may be seen as a weakness. This casts doubt on the validity of data that rely solely on questionnaire responses.

The experience of stress has been linked to the extent individuals perceive themselves to be in control of a particular situation and Miller (1980), on the basis of tightly controlled laboratory experiments, argues that to be 'in control' of a situation makes it more predictable and hence less stressful. The same notion of being 'in control' can be applied to the teachers in the classroom, even though the variables are more complex.

The relationship between the effects of stress (teacher burnout) and locus of control, the degree to which individuals see themselves as responsible for what happens to them, was explored by McIntyre (1984) He showed that, on his measures, the two areas were positively related at a statistically significant level. Stress was measured on subscales of the Maslach Burnout Inventory (1981) and locus of control on the Nowicki-Strickland Adult Scale (1973). However, since the variance common to the two measures was little more than 2 per cent this means that the relationship is of low predictive value.

Ways for teachers to reduce stress are described by Iwanicki (1983). He advises working directly on symptoms of stress using techniques such as relaxation, exercise, improved diet and not dwelling on problems. Improved communication and shared power in the organisation can also reduce stress and techniques to improve these are put forward. An integration of both approaches has been developed by Hall *et al.* (1984). In their in-service courses they invite members to take more responsibility for the means they use to relate to themselves and to others. Evaluation of these courses permitted inference of a relationship between control and stress which was made without the use of conventional questionnaire data

---

* Now at Centre for Social and Moral Education, University of Leicester

The courses offered 6 days of experiential learning sessions during a 7-day period. A small group experience (Rogers, 1971) took up about one-third of this time, and a range of more structured learning situations dealt with counselling and inter-personal skills, assertion training, non-verbal communication, values clarification and the management of stress. The course members were not told what to learn, but were invited to develop their awareness of what they did and to explore new ways of behaving. In the light of this increased awareness they were encouraged to re-evaluate their responsibility, especially where habit may have led them to misperceive it. In psychological terms, they were encouraged to develop a more internal locus of control where this seemed to them appropriate.

METHOD

Two 6-day courses, involving 90 working teachers, were evaluated using a range of instruments which were also intended to have a heuristic value for the course members. (For more complete details of the course and evaluation see Hall *et al.*, 1984.) For the purpose of this article, only that part of the evaluation which dealt with the relationship between stress and control will be considered. This includes a diary kept by course members of significant events:

1. at work, before the course started, involving pupils and colleagues (N = 90)
2. during the course, with other course-members and persons in their private lives (N = 67)
3. at work, after the course, involving pupils and colleagues (N = 34)

For each incident, they were asked to state:

1. action you took
2. outcome
3. possible alternative action you might have taken
4. after effects (stress hangover)

Between 10 and 15 weeks after the course, 42 of the 90 teachers involved were given a one hour interview. These were chosen in terms of geographical convenience from 78 course members who had agreed to be interviewed. Interviews were arranged either at the place of work of the member, at their home or in the School of Education.

Interviewees were asked 10 questions, each of which was followed up by the interviewer with a summary of the content and feeling of the response as he perceived it. This led in some cases to interviews which lasted for more than an hour and a half, and in others to interviews which lasted for less than half an hour, depending upon the amount of time the interviewee had available and the mutual interest generated in the process of the interview. The following 10 questions were asked:

1. What effect did the short course have on you?
2. Have you attended similar courses before?
3. Has any other course had a similar effect?
4. If you were to sum up what you learned on the course, what would it be?
5. Who or what would you say was responsible for what you learned on the course?
6. How has what you learned changed your behaviour?

7. Who or what within your work setting has enabled you to maintain the changes you wished to make?
8. Who or what has worked against change?
9. The indication from diary entries is that course members generally allowed themselves greater flexibility in their responses to other course members than they usually do with colleagues at work. Is this true for you? How do you account for that?
10. What habits of behaviour did you become aware of during the course?

RESULTS

In the total of 90 pre-course diaries, 327 incidents are recorded. Of these, 187 concern incidents with pupils and 140 with other staff members. These reports describe a range of incidents that teachers experience as stressful and show that habitual responses to stressful situations tend to be ineffective, unplanned and potentially damaging to personal relationships.

Punishment or the threat of punishment was involved in three quarters (74 per cent) of actions taken with pupils as a result of recorded incidents. One-third (33 per cent) of interactions with other staff members also involved punishment. But, whilst many forms of punishment are available in dealing with pupils — setting extra work, detention, verbal violence, etc. — the single formal sanction against another member of staff is to report to a senior colleague. Informal punishments are more diverse and include sarcasm, rudeness, 'lecturing' and refusing to co-operate.

The list of incidents which elicit the punishing response and consequent stress reveals a striking similarity between incidents with pupils and incidents with other members of staff. Table 1 shows the two lists in rank order.

Punishment, in these circumstances, is the automatic response used to re-assert control over disruptive, or potentially disruptive, situations. It is a grossly inefficient tactic. Some success in maintaining the smooth running of the institution may result, but the ensuing stress for the teacher is reported to be high. The legacy of the punishing approach for the person using it was listed as damaged interpersonal relations, alienation, anxiety and personal suffering. These reports, taken together with evidence from the literature, make it clear that keeping discipline is the part of the teacher's role which generates high levels of stress.

This is evident to most teachers, but it is suggested that narrow traditional interpretations of the discipline function is seen as an institutional demand. Lacking encouragement from inside or outside their profession, poorly equipped by training and with little time to search out alternatives for themselves, teachers overlook the possibility of exercising discipline in a more creative way.

The Nottingham courses sought to facilitate the use of more creative responses to the experience of stress and to the disruptive events seen to cause it. The course did this

TABLE 1

A COMPARISON OF THE MOST FREQUENTLY RECORDED INCIDENTS IN PRE-COURSE DIARIES

Incidents with pupils (N = 187)		Incidents with staff (N = 140)	
Disrupton of lessons	(29%)	Disruption of admin. procedure	(16%)
Aggressive behaviour between pupils	(18%)	Aggressive behaviour between staff	(13%)
Pupils refusal to work as directed	(10%)	'Unprofessional' behaviour of colleagues	(8%)
Truancy	(7%)	Covering for absent colleagues	(6%)
Pupils wearing wrong uniform, or inappropriate clothing	(5%)	Feeling at odds with the system	(4%)
Inappropriate talking in classroom	(5%)	Communication breakdown in staff meetings	(4%)

by examining experientially ways in which habitual responses could be set aside and experiments with new strategies substituted. Such experiments included the use of stress as a cue to begin behaving differently

All course members who were followed-up described the open, small group as both a strong and positive learning experience. It was perceived as an extremely difficult situation because of the lack of formal guidance about how to behave. All commented on the anxiety generated by being in a position where responsibility for their own learning was made clear to them, and acceptance of this was requested. They also commented that this raised level of arousal in unstructured circumstances tended to result in an examination of habitual behaviour and to experiment with new behaviour. Habitual behaviour is most effectively suppressed when it is immediately challenged and an alternative and rewarding behaviour is available. In the small group increased directness with other course members was reported to be just such a rewarding alternative.

The same is reported to be true back in the work situation where increased directness led to improved relationships and a lessening of stress and feelings of having left a situation half dealt with or dealt with only at an institutional and impersonal level. Forty of 42 interviewees included statements of this nature.

Without exception interviewees reported changed behaviour as a result of the 6-day course. This was either change at a covert or insight level:

> "What the course gave me, which considerable reading in the area had not, was a capacity to appreciate that we each have enormous potential to solve our own problems, ease our tensions, and resolve our conflicts given the opportunity to work with a skilful and sympathetic other."

or at the level of concrete behaviour:

> "One very apparent outcome is my newfound confidence in meetings at school where I am now able to express my point of view without the (at one time) red face and racing pulse."

or at both levels simultaneously:

> "At a meeting of teachers and advisors . I realised I was being put into a stressful situation .. channelled into situations I did not want for personal as well as professional reasons. I felt stressed, but put my point forward coolly and assertively rather than aggressively. The realisation that I was becoming stressed helped me to clarify my thoughts and concentrate on the matter in hand rather than concentrate on being stressed. As a result of the course I have become more aware of myself and my relationships with others, and I am developing more effective ways of communicating."

DISCUSSION

The data reported in this paper support the earlier research quoted at the beginning in finding that teaching is a stressful activity The diaries and interviews provide more open-ended data than the standardised questionnaires employed in that research An additional difference is that the data were collected during a course which encouraged members to be more open about their experience. In fact, the measures were seen as heuristic devices which would themselves affect pupil behaviour. Thus, no direct comparison with the research quoted can be made, but the course did appear to encourage a higher proportion of reports of stress induced by colleagues.

The main finding, however, is that an experiential in-service course of this nature enables teachers to change their behaviour in a way that reduces stress.

Without exception the in-course diaries report that the course itself was stressful, particularly the small group, providing the course members with an opportunity to work directly on issues related to stress.

Another main theme of the course involved taking responsibility for, and hence control of, more of one's own behaviour. Most course members reported that they had taken responsibility for changing their behaviour at work, so that it seems reasonable to link the process of taking responsibility for one's own behaviour to a reduction in stress.

At the end of his paper Smilansky (1984), drawing from work in family therapy, suggests that training institutions should assist teachers to reduce stress by becoming more open and presenting themselves more realistically, rather than cultivating a professional stereotype. The results from this study provide strong support for this hypothesis.

The skills training provided by the Nottingham courses is much looser than microteaching (Hargie, 1984) in that students decide for themselves what to learn after having been helped to categorise behaviour in relevant areas, and how their learning can be applied appropriately in the work setting. There are good reasons why this approach should be taken, since an in-service course can never provide an exact replica of future problems. Experiential learning enables the students to learn how to learn, allowing them to use their own authority rather than infantilising them with advice from the expert.

REFERENCES

Dunham, J (1976) Stress situations and responses. In National Association of Schoolmasters (Ed.), *Stress in Schools* National Association of Schoolmasters.

Dunham, J (1980). An exploratory comparative study of staff stress in English and German comprehensive schools *Educ Rev*, 32, 11-20

Fletcher, B C, and Payne, R L (1982). Levels of reported stressors and strains amongst schoolteachers some UK data *Educ. Rev*, 34, 267-278.

Hall, E, Woodhouse, D, and Wooster, A. D (1984) An evaluation of in-service courses in human relations *Br J. In-Service Educ*, 11 (1), 55-60.

Hargie, O D. W (1984) Training teachers in counselling skills the effects of microcounselling *Br J educ Psychol.*, 54, 214-220

Iwanicki, E F (1983). Towards understanding and alleviating teacher burnout *Theory into Practice*, 12, 27-32.

Kyriacou, C, and Sutcliffe, J (1978) Teacher stress prevalence, sources and symptoms *Br. J educ Psychol*, 48, 159-167

Maslach, C, and Jackson, S E (1981) *The Maslach Burnout Inventory* Palo Alto, Calif.: Consulting Psychologists Press

Miller, S. M. (1980) Why having control reduces stress· if I can stop the roller coaster, I don't want to get off. In Garber, J, and Seligman, M E. P. (Eds), *Human Helplessness Theory and Applications* New York Academic Press

McIntyre, T C (1984) The relationship between locus of control and teacher burnout. *Br J educ Psychol*, 54, 235-238

Nowicki, S., and Strickland, B R (1973) A locus of control scale for adults *J. consult clin Psychol*, 40, 148-154

Rogers, C R (1971) *Encounter Groups* London Penguin Press.

Smilansky, J. (1984). External and internal correlates of teachers' satisfaction and willingness to report stress *Br J educ Psychol.*, 54, 84-92.

(*Manuscript received* 10*th December*, 1984)

Br. J. educ. Psychol., 55, 124-137, 1985

RELATIONSHIPS BETWEEN SCHOOL MOTIVATION, APPROACHES TO STUDYING, AND ATTAINMENT, AMONG BRITISH AND HUNGARIAN ADOLESCENTS

By N. J. ENTWISTLE
(Department of Education, University of Edinburgh)

and B. KOZÉKI
(Institute of Psychology, Budapest)

Summary. Samples of 614 British and 579 Hungarian pupils aged 13-17 completed an inventory covering 20 sub-scales of school motivation and approaches to studying Indices of attainment in arts and science were also obtained. The factor structure of the inventory proved to be remarkably stable across nationality and sex, with school motivation accounting for two factors (the emotional and moral domains, and the cognitive domain) and approaches to studying dividing into three (meaning, reproducing, and organising/strategic) Correlations of the scales with attainment were obtained, along with multiple regression analyses. The overall inventory provides valuable information to teachers in the form of not just a single index of school motivation but also a profile of scores indicating different types of motive and ways of tackling school work.

INTRODUCTION

Kozéki (1984) has described a model of school motivation which places equal emphasis on three separate domains — affective, cognitive, and moral. It seeks to explain pupils' school behaviour and attainment in terms of an interaction between pupils' relationships with parents, teachers or peers (the affective domain) and their developing demands for independence, competence, and interest within their school work (the cognitive domain). The moral domain contains, in effect, one type of outcome of earlier interactions — the growth of trust, compliance or responsibility. A summary of the nine dimensions is shown in Figure 1.

The initial description of these nine motives was based partly on interviews and partly on experimental work which related actions to the consequences anticipated by the pupils (Kozéki, 1975, 1980). Subsequently a *School Motivation Inventory* was developed which produced six factors, two in each domain. This early work was carried out entirely in socialist countries in Central Europe, but recently a comparative study

FIGURE 1

Sub-scales and Factors of School Motivation
(adapted from Kozéki and Entwistle, 1984, page 308)

	DOMAINS		
	Affective (A)	*Cognitive* (B)	*Moral* (C)
	Warmth (Parents)	Independence	Trust
MOTIVES	Identification (Teachers)	Competence	Compliance
	Sociability (Peers)	Interest	Responsibility
	Perceived pressure from Adults (−ve)		

was reported (Kozéki and Entwistle, 1984) which confirmed the existence of between five and eight oblique factors within the 60 items of the inventory. But as the scales and factors all had fairly high intercorrelations, a substantial general factor of school motivation could also be inferred, and was indeed intended. This previous work thus indicated the conceptual value of identifying up to nine motives within three domains when describing school motivation, and yet the empirical studies indicated substantial overlap between the various aspects of motivation. However, strong interrelationships should not be taken to imply psychological identity. It is equally appropriate, for different purposes, to stress empirical connections or conceptual separations.

The relationship between academic motivation and scholastic success is well established, but recently it has been suggested (Entwistle *et al.*, 1974; Biggs, 1978) that differing types of motivation are associated with characteristically different ways of studying. Initially four distinct types of student were identified by cluster analyses from scales of personality, motivation, study methods, and attitudes (Entwistle and Wilson, 1977): well-organised and competitive; anxious, syllabus-bound, working long hours, hard-working, syllabus-free, with aesthetic interests and radical attitudes; and academically unmotivated with poor study methods.

FIGURE 2

SUB-SCALES OF APPROACHES TO STUDYING
(adapted from Entwistle and Ramsden, 1983, page 42)

Orientation	*Approach or Style*	*Motivation*
Meaning	Deep Approach Holist Style	Intrinsic
Reproducing	Surface Approach Serialist Style	Fear of Failure or Instrumental
Achieving	Strategic Approach Conscientiousness	Hope for Success

Subsequent research has concentrated on operationalising the concepts of Marton (1984) and Pask (1976) within an *Approaches to Studying Inventory* (Entwistle and Ramsden, 1983). Factor analyses of this inventory established the existence of four main *study orientations*: meaning, reproducing, achieving or strategic, and non-academic. The term 'orientation' was used to indicate a combination of approaches to studying and styles of learning which is relatively stable across different educational tasks. These orientations are also associated with characteristic forms of motivation, as shown in Figure 2. Although three of the orientations were consistent across all six subject areas investigated, the strategic orientation was less clear-cut. It appeared most characteristically as a combination of strategic approach ('cue-seeking' in Miller and Parlett's, 1974, terms) and the competitive drive of achievement motivation, making it reminiscent of the first of the clusters mentioned above. But another factor was also produced which linked achievement motivation with either intrinsic or extrinsic motivation. Parallels between the descriptions of Kozéki's motivation and Entwistle's student types and approaches to studying have already been discussed (Kozéki and Entwistle, 1983). This article reports the empirical relationships between the various facets of school motivation and the sub-scales of a version of the *Approaches to Studying Inventory* designed for use with secondary school pupils.

METHOD

Inventory of 'School and School Work'

The English version of the *School Motivation Inventory* was developed by correspondence between the authors. As a result some items failed to carry the full meaning of the Hungarian originals, and the English scales consequently had lower internal reliabilities. Subsequent discussions of the item analyses from the first study allowed the Hungarian and English versions to be closely matched for meaning. Besides identifying the main factors, the item analyses also suggested the possible importance of perceived parental and teacher pressure on pupils' attitudes and motivation, and a scale of 'perceived pressure from adults' was added to the original nine motives. The *Approaches to Studying Inventory* was also discussed and nine of the 16 scales were selected, with wordings adjusted to suit the school context in both Britain and Hungary. The intention was to retain, as far as possible, the main defining scales of the meaning, reproducing, and strategic orientations. However, as pupils have little opportunity for strategic studying, in the sense of 'cue-seeking', the strategic approach was defined, instead, as a highly organised way of tackling school work with an eye to good attainment. The sub-components of deep approach (relating ideas and use of evidence) were omitted, as were the learning style pathologies (globetrotting and improvidence). Instead of 'negative attitudes', a positive scale of 'conscientiousness' was used. An indication of the meaning of the 20 scales can be obtained from the defining items shown in Table 1. (The percentage responses will be discussed in a later section.)

TABLE 1

DEFINING* ITEMS OF THE SUB-SCALES AND PERCENTAGE AGREEMENT† BY NATIONALITY AND SEX

Sub-scale	Item	Defining Item	British		Hungarian	
			Boys (285)	Girls (329)	Boys (300)	Girls (279)
Warmth	(A31)	My parents are always helpful and encouraging about my school work	80	84	86	86
Identification	(A52)	Most teachers try hard to help all the pupils	**75**	**76**	53	55
Sociability	(A13)	I enjoy helping other pupils with their school work	52	67	**80**	**92**
Independence	(A44)	I always prefer to work things out for myself	69	63	**81**	**77**
Competence	(A35)	The things we learn at school are not of any real use to me (—)	26	18	22	20
Interest	(A46)	There are a lot of lessons which I find exciting and challenging	64	66	59	60
Trust	(A27)	If I'm given something to do, I always try to do it as well as possible	80	84	**88**	**94**
Compliance	(A28)	School rules are sensible I always try to follow them	47	**58**	53	**57**
Responsibility	(A49)	I am ready to take responsibility for all my actions, no matter what	75	67	77	80
Adult Pressure	(A60)	Adults always seem to be expecting too much of young people	38	40	33	39
Deep Approach	(B51)	In trying to understand new ideas, I often try to relate them to real life situations	63	72	**77**	**78**
Holist Style	(B22)	I enjoy doing things where I can use my imagination or my own ideas	82	81	**91**	**90**
Intrinsic Motiv	(B53)	I get very enthusiastic about some of my school work	56	52	47	56
Surface Approach	(B4)	I find I have to rely on memorising a good deal of what we have to learn	**66**	**65**	49	39
Serialist Style	(B5)	I prefer to tackle each part of a topic or problem in order, working through them one step at a time	**76**	**86**	39	51
Fear of Failure	(B27)	I am always worrying that I will get behind with my work	**50**	**52**	43	38
Instrumental Motiv	(B6)	I suppose I'm more interested in the qualifications I'll get than in the subjects I'm taking	**54**	**49**	**43**	27
Strategic Approach	(B58)	I plan my working time carefully, to make the most of it	40	36	**50**	**49**
Conscientiousness	(B30)	I don't mind working long hours to complete my work satisfactorily	60	62	60	65
Hope for Success	(B29)	It's important to me to do things better than other pupils, if I possibly can	**57**	43	**53**	43

* In terms of correlation with corrected sub-scale total † 'Strongly agree' and 'agree' responses combined

Bold figures draw attention to groups with the highest levels of agreement.

Other information collected included sex, preference for arts or science subjects (or both or neither), and indices of attainment. These indices were on a 50-10* scale based in the Hungarian sample on the average of teachers' ratings on, respectively, arts, science and overall attainment, and in the British schools on either self-ratings or, where available, the average O-level grades converted on to a similar 50-10 scale. The subjectivity of these indices for most of the sample is admitted, but the measures were the only comparable ones available. As it was intended that the inventory would be completed in 30-40 minutes (in either one or two sittings) it was necessary to restrict the overall length of the inventory to 120 items. The length of each sub-scale was thus restricted to 6 items, with consequent limitations on reliability.

Pilot work indicated that pupils generally maintained interest in the inventory for its full length, as the items were seen as personally relevant and had been designed to be colloquial. Instructions indicated that the inventory contained comments made by pupils about their school and school work, and asked "To what extent do you agree or disagree with their comments?" (see Kozéki and Entwistle, 1984, p. 309, for further details). Responses on an 'a'-'e' scale were coded 5-1, giving a maximum of 30 for each sub-scale. Estimates of internal consistency (Cronbach alpha) and test-retest reliability are reported later (Table 3). The validity of the school motivation sub-scales was discussed previously (Entwistle and Kozéki, 1984, pp. 309 and 317), while similar evidence of the internal validity (Entwistle and Ramsden, 1983, Chapter 4; Watkins, 1983) and external validity (Watkins, 1983, p. 34) has been provided for the student version of the *Approaches to Studying Inventory*. Watkins' work is particularly valuable as it involved using the inventory in a different educational system (Australian). He reproduced the main factors in the inventory and showed 78 per cent agreement between inventory scores (dichotomised) and interview ratings of approach to studying.

Sample

As the approaches to studying had been identified previously only in samples of students in higher education, it was decided to collect data mainly from older pupils with higher levels of academic attainment, but to include also a comparison group of younger pupils to see whether these scales were, in fact, still effective. The sample was thus made up of pupils aged 15-17 and 13-14 with an intended bias of about 2:1 in terms of the proportion of older pupils.

TABLE 2

COMPOSITION OF THE SAMPLES

Age-groups	Britain			Hungary			Overall Total
	Boys	Girls	Total	Boys	Girls	Total	
16/17	161	145	306	102	83	185	491
15/16	48	99	147	103	97	200	347
13/15	76	85	161	95	99	194	355
Total	285	329	614	300	279	579	1193

In Britain, four schools were chosen which retained a high proportion of pupils beyond age 16. Three schools were in Scotland and one in Northern England; two in a city and two in small towns. In Hungary a similar sample was obtained by taking four academic gymnasia (which prepare pupils for university beyond age 15). Most of the younger children came from the 'general' schools (up to age 15) of two of the gymnasia, and so were of above average ability. Only one small 'general' school contained pupils of lower ability. The composition of the sample in terms of age, sex, and nationality is shown in Table 2.

---

* In Hungary, teachers' marks are always given on a 5-1 scale. Using these, averages could be obtained across the various subjects correct to one decimal place, then multiplied by 10 to remove the decimal place.

It was possible to give the inventory on a second occasion to 391 Hungarian pupils aged 13/14 and 15/16 after an interval of about three weeks.

The main analyses were designed to describe salient characteristics of the sub-groups, and to investigate the inter-relationships between school motivation, approaches to studying and school attainment, separately in Britain and Hungary.

RESULTS

Characteristics of the sub-groups

Table 3 presents the estimates of consistency and reliability of the 20 sub-scales, along with the means on those scales by nationality, sex, and subject preference (showing only those who made a definite choice of either arts or science). Statistical significance of differences in mean scores was tested by 't' tests and is indicated, at the 0·05 level, by asterisks in Table 3.

The reliability of the sub-scales suggested weakness in some of the sub-scales. 'Sociability', 'independence', 'holist', 'serialist', and 'hope for success' were consistently low in consistency indices in both samples, and most dropped below 0·65 in test-retest reliability as well. The other scales showed a level of reliability acceptable for scales of this length, although future developments of the inventory will increase reliability by combining similar sub-scales (e.g., trust and responsibility).

The main differences in mean scores were between the two nationalities, with the Hungarians having higher scores in all the cognitive and moral motives, and all the sub-scales of meaning orientation. British pupils had higher scores on identification (relationships with teachers), hope for success, and all the sub-scales of the reproducing orientation (including serialist style).

Consistent sex differences showed girls to have higher scores on warmth (relationships with parents), sociability (with peers), and all three sub-scales in the moral domain Hungarian girls had higher scores on deep approach and lower scores on surface approach; British girls showed higher levels of fear of failure than any of the other sub-groups. Boys consistently scored more highly on instrumental motivation and hope for success.

The differences between pupils indicating preferences for arts or science subjects showed fewer instances of statistical significance (owing to the smaller sample sizes). There were, however, fairly consistent tendencies for those preferring science to have higher cognitive and moral motives, and higher scores on serialist style (with the exception of the small group of Hungarian girls with scientific preferences).

The interpretation of the differences in mean scores can be facilitated by drawing on the percentage responses to the defining items of the sub-scales, presented in Table 1. It appears that pupils in the British sample showed more respect for their teachers (A52 — 76/54%), but were more prone to rely on memorisation (B4 — 66/44%) and to adopt a rather rigid procedural approach to learning (B5 — 81/45%). Pupils in the Hungarian sample were more likely to try to relate school learning to real life (B51 — 78/68%), to plan their time carefully (B58 — 50/38%), and to enjoy helping other pupils (A13 — 86/60%).

TABLE 3

MEAN SCORES AND RELIABILITY ESTIMATES OF SUB-SCALES — BY NATIONALITY, SEX, AND SUBJECT PREFERENCE

Sub-scales	British								Hungarian							Reliability		
	Boys				Girls				Boys				Girls			Internal		T/Rt
(N)	Arts (78)	Sci (128)	Tot (285)	B/G	Arts (131)	Sci (103)	Tot (329)	B/H†	Arts (78)	Sci (111)	Tot (300)	B/G	Arts (145)	Sci (30)	Tot (279)	Brit.	Hung.	Hung.
School Motivation																		
Warmth	16·4*	17·8	17·4	*	18·9	18·7	18·6		18·7	18·0	18·1	*	18·8	19·4	19·0	0·74	0·67	0·74
Identification	13·5	14·4	14·1		14·5	14·7	14·5	*B	12·3*	14·1	13·4		13·4	14·6	14·0	0·68	0·71	0·78
Sociability	14·3	14·9	14·6	*	16·7	16·4	16·4	*H	16·3	17·2	16·8	*	17·9	17·8	18·0	0·41	0·50	0·63
Independence	15·7	16·6	16·3		16·3	16·9	16·3	*H	16·9	17·3	17·0		16·7	18·2	16·8	0·51	0·52	0·65
Competence	14·4*	15·7	15·1		14·8	15·7	15·2	*H	15·9	16·8	16·2		15·6*	17·3	16·2	0·57	0·57	0·68
Interest	12·4	13·1	12·8		13·0	13·3	13·1	*H	13·7*	16·1	15·1		15·4	16·4	15·7	0·70	0·63	0·72
Trust	15·5	16·0	16·0	*	16·8	17·2	16·8	*H	16·9*	18·0	17·3	*	18·8	18·6	18·9	0·56	0·62	0·76
Compliance	13·8*	15·1	14·7	*	16·0	16·9	16·2	*H	15·1	16·2	15·6	*	16·4	17·5	16·8	0·68	0·62	0·77
Responsibility	16·2	16·9	16·9	*	17·8	18·0	17·8	*H	18·0	18·0	17·9	*	18·6	19·7	18·7	0·51	0·60	0·72
Adult Pressure	16·3*	14·9	15·4		14·5	15·0	14·8	*B	13·5	13·8	13·8		14·1*	12·0	13·6	0·72	0·67	0·73
Approaches to Studying																		
Deep Approach	15·5	16·5	15·9		15·7	15·8	15·6	*H	16·5	17·1	16·8	*	17·5	17·9	17·7	0·65	0·64	0·72
Holist Style	16·6	17·3	17·2		17·7	16·9	17·4	*H	18·7	18·4	18·3		18·9	18·6	18·8	0·49	0·51	0·67
Intrinsic Motivation	12·8	12·8	12·6		12·6	12·5	12·5	*H	14·3	15·3	14·6		15·1	16·3	15·3	0·79	0·75	0·74
Surface Approach	12·6	12·5	12·8		12·2	13·1	12·6	*B	9·2*	11·1	10·5	*	8·9	8·8	9·4	0·53	0·61	0·73
Serialist Style	14·5*	15·6	15·3		15·6	16·1	15·6	*B	12·3*	13·4	12·9		12·7	11·3	12·7	0·45	0·32	0·63
Fear of Failure	11·6	10·5	11·1	*	13·0	13·3	13·0	*B	10·6*	11·9	11·7		12·1	10·6	11·9	0·64	0·70	0·77
Instrumental Motivation	12·5	12·4	12·8	*	11·3	11·7	11·6	*B	10·2	10·5	10·7	*	9·8	8·6	9·2	0·54	0·66	0·76
Strategic Approach	12·4	12·8	12·7		12·5	12·6	12·6	*H	13·7	14·1	14·0		13·9	14·8	14·6	0·64	0·70	0·76
Conscientiousness	14·4	15·0	14·9		14·9	15·0	14·8		14·7	15·1	14·8		14·7	15·5	15·5	0·74	0·76	0·73
Hope for Success	13·1*	14·9	14·2	*	12·6	13·3	12·6	*B	10·0	10·7	10·6	*	9·0	10·0	9·4	0·50	0·49	0·69

* Indicates statistical significance at 0·05 level or better

† Statistical significance of differences between British and Hungarian samples, letter indicates which mean was higher

Correlations with school attainment

Table 4 reports first-order correlations between each of the 20 sub-scales and the indices of arts and science attainment for British and Hungarian boys and girls, separately.

TABLE 4

CORRELATIONS OF SUB-SCALES WITH ARTS AND SCIENCE ATTAINMENT BY NATIONALITY AND SEX

	British				Hungarian			
	Boys (285)		Girls (329)		Boys (300)		Girls (279)	
Sub-scales	Arts	Sci	Arts	Sci	Arts	Sci	Arts	Sci
School Motivation								
Warmth	0·14	0·16	0·30	0·30	0·12	0·06	0·16	0·25
Identification	0·22	0·14	0·20	0·20	0·06	0·07	0·08	0·27
Sociability	0·05	0·14	0·19	0·14	0·07	0·08	0·10	0·20
Independence	0·24	0·20	0·33	0·29	0·21	0·19	0·24	0·20
Competence	0·22	0·28	0·24	0·26	0·17	0·20	0·12	0·25
Interest	0·11	0·18	0·19	0·26	0·09	0·13	0·09	0·15
Trust	0·14	0·10	0·20	0·21	0·02	0·03	0·16	0·21
Compliance	0·19	0·20	0·31	0·33	0·08	0·09	0·17	0·24
Responsibility	0·24	0·25	0·17	0·16	0·09	0·09	0·17	0·16
Adult Pressure	−0·34	−0·26	−0·28	−0·27	−0·32	−0·25	−0·38	−0·38
Approaches to Studying								
Deep Approach	0·11	0·13	0·27	0·23	0·22	0·20	0·17	0·17
Holist Style	−0·03	0·09	0·13	0·07	0·16	0·11	0·08	0·03
Intrinsic Motivation	0·10	0·15	0·23	0·26	0·16	0·14	0·03	0·12
Surface Approach	−0·17	−0·14	−0·23	−0·13	−0·25	−0·16	−0·38	−0·20
Serialist Style	0·02	0·10	0·07	0·12	−0·15	−0·08	−0·22	−0·16
Fear of Failure	−0·25	−0·34	−0·18	−0·19	−0·21	−0·21	−0·32	−0·24
Instrumental Motivation	−0·10	−0·14	−0·34	−0·23	−0·21	−0·16	−0·38	−0·36
Strategic Approach	0·15	0·09	0·18	0·20	0·12	0·17	0·09	0·20
Conscientiousness	0·09	0·12	0·19	0·20	0·12	0·08	0·12	0·25
Hope for Success	0·11	0·23	0·22	0·20	0·09	0·07	−0·13	−0·02

Correlations >0·11 are significant at 0·05 level, >0·15 at 0·01 level

The most consistent correlations indicate perceived pressure from adults being associated with low levels of attainment, and positive relationships of attainment with independence, competence and deep approach. In Britain, identification with teachers, compliance and warmth (both among girls), low fear of failure (among boys), and low scores on surface approach and instrumental motivation (both among girls) show significant relationships with the indices of attainment. In Hungary, the strongest relationships with attainment come from the variables which have negative correlations, and the closer relationships are found among the girls. One interesting difference between Britain and Hungary is in the competitive dimension, 'hope for success', and the serialist style which show consistent (although not high) positive correlations with attainment in the British sample and negative (or near zero) correlations in the Hungarian sample.

Differences in the relationships with arts and science attainment can be seen in identification (high with arts in British boys and with science in Hungarian girls), and competence (higher in science but mainly for the Hungarian girls). A distinction can also be seen between relationships in surface approach (with higher negative correlations in arts) and serialistic style which is marginally more favourable to science attainment in Britain.

Inspection of the 20 × 20 sub-scale inter-correlation matrix for both samples (not reported here) showed rather high inter-correlations among the school motivation sub-scales (particularly within each domain) and also high correlations between intrinsic motivation and the cognitive motivational domain (0·64 in Britain and 0·55 in Hungary), and between conscientiousness and the moral domain (0·65 and 0·64). Using this information, together with the first-order correlations with overall attainment, 10 variables were selected to carry out regression analyses on total attainment for British and Hungarian boys and girls. The variables which significantly increased the multiple correlations with the criterion are shown in Table 5, in the order in which they were included.

TABLE 5

REGRESSION ANALYSES ON OVERALL ATTAINMENT: ORDER IN WHICH PREDICTIVE VARIABLES INTRODUCED

	British		Hungarian	
Predictive variables	Boys	Girls	Boys	Girls
Affective Domain (Warm + Ident + Soc)				4
Cognitive Domain (Indep + Compt + Int)	(5)	1		(6)
Moral Domain (Trust + Cmpl + Resp)	2	4	(6)	
Adult Pressure	−3		−1	−1
Deep Approach				
Surface Approach			−2	−2
Strategic Approach			(5)	
Fear of Failure	−1	−2	−3	−3
Instrumental Motivation		(−5)		−5
Hope for Success	4	3	4	
Multiple Correlation	0·53	0·50	0·42	0·50

( ) indicates that the introduction of this variable increases the multiple regression at a level just less than statistical significance. The negative sign indicates the direction of the beta coefficient

The variables picked out simply reinforce the impression created by the first-order correlations of the predictive strength of the 'negative' variables, but it is interesting to see the sex difference on 'hope for success' (included only for the boys) and the affective domain (included only for the Hungarian girls). The highest multiple correlation is 0·53, but this value is, no doubt, affected by the limited reliability of the attainment indices. and in any case can be considered high for this type of inventory.

It is perhaps interesting to note that adult pressure, although having the highest simple correlations with attainment, was entirely omitted from the prediction of attainment among British girls. Examining the partial correlation coefficient showed that adult pressure was a good predictive variable only because of its overlap with other measures. Including adult pressure resulted in the suppression of variables with greater explanatory potential, such as cognitive motives. Overall, the lack of stability across samples suggests that the predictive variables are too closely inter-related to produce consistent beta-weights in the regression equations, and only the estimate of multiple correlation can be treated with confidence.

TABLE 6

FIVE-FACTOR SOLUTIONS FOR BRITISH AND HUNGARIAN SAMPLES

Sub-scales/Factor	British Schools (N=614)					Hungarian Schools (N=579)					Scale
	I	V	IV	II	III	I	V	IV	II	III	
School Motivation											
Warmth	71	37			31	62	42				Warm
Identification	71	32			44	68	53				Ident
Sociability	52					53	34				Soc
Independence	59	43				56	56	36	−48		Indep
Competence	58	57			66	58	52	32		43	Compt
Interest	52	40			83	55	73	36		56	Int
Trust	60	57			51	67	77				Tru
Compliance	68	54			59	71	59				Compl
Responsibility	61	44			33	63	35				Resp
Adult Pressure	−70			37	−31	−65	−49		45		Press
Approaches to Studying											
Deep Approach	35	52	57		50			70		44	Deep
Holist Style			65					74			Hol
Intrinsic Motivation	34	42	43		75	32	48	47		61	Intri
Surface Approach				62				−34	77		Surf
Serialist Style	34	43		46					54		Seri
Fear of Failure				58					65		Fear
Instrumental Motivation				59	31	−32			63		Inst
Strategic Approach	36	65			42	39	69			35	Strat
Conscientiousness	51	79			54	45	84			34	Consc
Hope for Success		46								56	Succ
Percentage variance associated with each unrotated factor	34·0	5 3	5 5 Total 63 4%	11 1	7 5	32 1	4·7	5·4 Total 65 4%	14·1	9 1	

(Decimal points and loadings less than 0·30 omitted)

Factor analyses

In a situation where the variables in a factor analysis are drawn from domains that are known to be overlapping, there may be chance variations in the factor structure from sample to sample. It is thus important to investigate the stability of the factor structure across sub-groups. The main analyses of the 20 sub-scales were thus repeated for boys, and girls, and for pupils aged 15/17 and 13/14 in each country. Additional analyses were also carried out with the two independent attainment variables added, and for each section of the inventory separately.

The SCSS program (Nie *et al.*, 1981) was used, selecting the principal axes analysis, followed by oblique rotation (with delta set at zero). An eigen value of unity was used as the main criterion for extracting factors, but as the eigen values varied somewhat from group to group the factor-solution reported here is the one which was associated with eigen values close to unity in a majority of the analyses of that set of variables. Table 6 presents the factor analyses of the 20 variables for British and Hungarian total samples separately, while Table 7 shows the inter-correlations between the oblique factors. These five-factor solutions extracted 63·4 per cent and 65·4 per cent of the variance respectively.

TABLE 7

INTERCORRELATION OF OBLIQUE FACTORS

	British				
	I	II	III	IV	V
I		−0·04	0·44	0·15	0·43
II	−0·15		0·05	0·00	0·07
III	0·10	0·04		0·19	0·39
IV	0·21	−0·22	0·23		0·25
V	0·58	0·07	0·28	0·25	
	Hungarian				

Factors

I Warmth and Moral Domains
II Reproducing Orientation
III Interest, Competence & Conscientiousness
IV Meaning Orientation
V Moral and Cognitive Domains

In these analyses the three domains of school motivation were not clearly separated. Two overlapping factors (correlating 0·44 and 0·58 respectively in the British and Hungarian samples) loaded heavily on most of the nine motives. The distinction between the two factors was that the first factor emphasised relationships with parents and teachers (linked to the moral domain, and particularly with compliance), while the second factor had its highest loadings on 'conscientiousness' and 'strategic approach', along with 'trust' and either 'competence' or 'interest'. In separate factor analyses of the school motivation sub-scales, the pattern matrix showed the three domains more clearly, but with still a tendency for the affective and moral domains to merge.

The remaining three factors showed a majority of high loadings on the 'approaches to studying' sub-scales. Factors IV and II clearly represented meaning and reproducing orientations respectively, with their associated motivational components. It is interesting to note that the reproducing orientation was also associated with 'perceived pressure from adults' in both samples, and with lower levels of 'independence' in the Hungarian sample.

Factor III represents what was termed the 'strategic or achieving' orientation among students. Here it represents an organised and conscientious approach to studying associated more with interest than with competitiveness. In separate analyses of the 'approaches to studying' sub-scales, reproducing orientation remained consistently distinct, but meaning orientation tended to overlap with the organised and conscientious approach.

Among the younger pupils in both samples a separate factor of 'holistic style' emerged, associated with 'competence' and 'organised approach', but with a negative loading on 'hope for success'. In this age-group 'meaning orientation' could not be separated from the two school motivation factors. Instead a factor of 'conscientiousness and compliance' was formed linked to 'trust', and to 'identification' (with teachers) in the Hungarian sample only.

Only the sub-group analyses with younger children have been mentioned here, because the factor analyses for boys, girls and 15-17-year-olds were very similar to those for each main sample.

The inter-correlations of the oblique factors, shown in Table 7, indicate the substantial overlap between the factors describing the school motivation sub-scales. It is interesting to note that, in the British sample, there was no relationship between the factors describing meaning and reproducing orientations, while among Hungarian pupils a small negative correlation was obtained. The other substantial difference in correlation was between Factors I and III, but the smaller correlation in the Hungarian sample indicates only that the school motivation scales are spread over three factors in the British sample (I, III and V), rather than two (I and V).

Factor analyses were also carried out in which arts and science attainment indices were added to the twenty sub-scales. The factor structure remained very similar, with school attainment producing a tightly defined additional factor which also showed loadings on a lack of adult pressure in both samples, and additional loadings on independence and low fear of failure in the British sample.

DISCUSSION

The importance of cross-cultural studies in psychology has been clearly recognised in the field of cognitive development, but in other areas of research rather few such studies have been reported. However, Brislin (1983) has argued that such cross-cultural research is essential in separating the more fundamental, psychological 'universals' of human behaviour from those which have their origins more in the particular social setting. If it is possible to demonstrate the existence of very similar patterns of relationships between variables in quite different social settings, then the evidence for the psychological reality and explanatory potential of those variables will be substantially strengthened. It was with this idea in mind that the current collaborative work was undertaken.

In our first study (Kozéki and Entwistle, 1984) it was shown that the *School Motivation Inventory*, developed in Central Europe, worked effectively in British schools, and that factors with a similar meaning could be identified in both Britain and Hungary. The intention of the present study was first to carry out a similar test of the school version of the *Approaches to Studying Inventory*, and then to see whether similar patterns of relationships could be discovered among the 20 sub-scales of the combined *School and School Work Inventory*.

The corrected item/sub-scale correlations (not reported here) show that the same group of items define equivalent scales in Britain and Hungary. There was some overlap between four sub-scales. 'Interest' and 'intrinsic motivation' were essentially the same, while 'trust' showed considerable overlap with 'conscientiousness'. Otherwise, although there were substantial correlations between sub-scale totals, individual items could be placed with confidence in their separate sub-scales in both countries. It is therefore clear that both sets of sub-scales can be said to work effectively in Britain and Hungary.

The similarity of the factor structure from the 20 sub-scales is quite remarkable, given the contrasting social and educational systems experienced by the pupils. The factors of 'school motivation', linking the three domains identified by Kozéki, 'meaning orientation', and 'reproducing orientation', are convincingly consistent between the main samples, and across sub-group analyses by sex and age. 'Strategic orientation', as in the student version (Entwistle and Ramsden, 1983), was less consistent. In the school version the meaning of 'strategic approach' had been changed. It no longer implied 'cue-seeking', but a strategic use of time and effort in carrying out school work. Its close association with conscientiousness is thus understandable, and these two variables defined the third main grouping of study orientations. Again, as in the research on students, rather separate factors defining academic motivation (interest, competence, intrinsic motivation, sometimes linked to hope for success) and a holistic style emerged from one or more of the sub-group analyses.

The three main study orientations present in the factor analyses provide good support for the model shown earlier as Figure 2. It appears that this structure applies equally to studying at school and in higher education, although with a different definition of 'strategic studying'. It could be argued that in schools a positive form of the 'non-academic orientation' has been merged with the previous strategic orientation to form the new strategic factor.

The sub-scales describing school motivation, however, do not form into the clearly distinct domains shown in Figure 1. Two factors repeatedly covered most of the sub-scales, showing only a different balance in the positive loadings. Where separation was possible in analysing these 10 scales alone, the cognitive domain emerged as a separate factor while the affective and moral domain remained together There is something of a paradox here. In the factor analyses of items reported previously (Kozéki and Entwistle, 1984) there seemed to be good justification for identifying between five and eight distinct dimensions — and the clear pattern of item/sub-scale correlations strongly supports their separate existence. But these factors, and the sub-scales, do have substantial inter-correlations as befits the component parts of an overall school motivation scale. Thus when the scales are intercorrelated and, in particular, when they are analysed alongside scales from two rather different domains (meaning and reproducing orientation), the unity of the school motivation scale is emphasised

Kozéki (1985) has argued that the three domains are complementary facets of school motivation Very few pupils have positive motivation in all three areas, but strong motives in only one domain would imply a potentially serious imbalance For example, over-reliance on emotional and social rewards may carry with it immature dependence: balanced by moral motives, greater self-reliance is implied. Similarly, seeking rewards exclusively in the cognitive domain may be associated with aggressive competitiveness unless balanced by either moral or affective motives. The factor analyses imply that, empirically, there is a general tendency for complementarity, a balancing of motives, particularly between the affective and moral domains. Complete separation of the domains could not be achieved, if there is also to be balancing between motives.

The utility of the individual sub-scales of school motivation was apparent in the analyses of means and in the correlations with school attainment. In analyses by sex or by subject area one or other sub-scales showed interesting, and plausible, patterns of relationships which differed from other motivational sub-scales. While caution has always to be observed in interpreting scores derived from self-report inventories, particularly between versions in different languages, the care that was taken in producing this version of the inventory allows greater confidence to be placed in these findings than in the results reported earlier.

One of the most noticeable differences in mean scores is found in relation to reproducing and meaning orientations. In the British schools there are much higher scores on surface approach and serialist style, together with instrumental motivation, and equivalently lower scores on deep approach (although the holist style is only slightly lower). In the later stages of secondary school the effect of the external examinations at 16, 17 (in Scotland) and 18, must be seen as the main reason for such high scores on reproducing. The syllabuses and examinations at O-level (or O-grade) do put a very heavy emphasis on the correct reproduction of information, and this inevitably has an effect on both the way teachers present the syllabus and the way children learn it Teachers have repeatedly complained about the over-loading of syllabuses (particularly in science) and the constraints it places on their teaching methods. Our inventory results, in comparison with Hungary, suggest that British pupils are being prevented from expressing their ideas, and exploring relationships between ideas — both of which are an essential part of developing personal understanding, whether in the arts or the sciences. It is tempting to suggest an explanation of the higher British scores on instrumental motivation in terms of the current levels of youth unemployment But without equivalent information for Britain from a period of lower unemployment any such interpretation must be viewed as speculative.

In the Hungarian schools, the high scores on deep approach and holist style are paralleled by low scores on surface approach and serialist style. At first sight this may seem to be an ideal combination, but the low scores on serialist style actually suggest a possible problem. Understanding depends on both comprehension and operation learning; the grasping of relationships has to be supported by an appropriate use of evidence and detail. The pattern of scores in the Hungarian schools suggest a tendency towards 'globetrotting' — the failure to support imaginative thinking by evidence and logical thinking From a knowledge of teaching methods in Hungary, this occurrence of rather unbalanced comprehension learning could have been anticipated. There has, in recent years, been a strong reaction against the previous emphasis on rote learning to the extent that both teachers and pupils have come to think of this method of learning as being unacceptable. In an attempt to foster creativity, as opposed to reproductive thinking, the emphasis both in lessons and text-books, has been on helping pupils to think about relationships. But the lack of emphasis on factual knowledge or operation learning is creating a problem which is currently being recognised. The pupils seem to lack an adequate knowledge-base from which to develop their ideas.

Of course, part of the explanation of the national differences may lie in response sets, but as the pattern is not uniform, this seems unlikely to be the main explanation. Thus our evidence would support the concern being expressed in Hungary about the over-emphasis on imaginative thinking in education, and in Britain on memorisation and surface approaches to learning. It seems that in both countries the emphasis on teaching is somewhat unbalanced, but in opposite directions. So each country could benefit from looking carefully at the teaching methods used in the other to see how a better balance could be achieved.

The prediction of school attainment from the combined inventory using multiple correlations was around 0·50 (or 25 per cent). In considering this level of prediction, it is important to recognise that neither the *School Motivation Inventory* nor the *Approaches to Studying Inventory* was designed specifically to predict school attainment. In fact, both were intended to move research away from this earlier preoccupation of educational psychologists. The qualitatively distinct approaches to studying, identified in students by Marton (1984) and more recently in secondary pupils by Selmes (1985), reflect the different intentions of students. Not all students, or pupils, see high academic attainment as their main goal. It is therefore unfortunate that the success of a pupil comes to be judged solely in terms of examination performance.

Similarly, the *School Motivation Inventory* was designed to break away from the previous tendency to investigate only a cognitive form of motivation (academic achievement motivation). Scores on the sub-scales of the *School Motivation Inventory* cover three domains and are intended to remind teachers and parents of their crucial contribution to the all-round development of children — affective, cognitive, and moral. Education in both countries has become excessively cognitive in its emphasis.

The inventory helps teachers to identify unbalanced motivational styles among pupils — and balance is as important as level of motivation in ensuring the development of the full potential of the individual pupil. Furthermore, the inventory provides information about ineffective approaches to studying which might require remedial action. Thus the combined inventory is intended to provide for teachers a wide-ranging instrument with which to explore pupils' motivational and learning styles and so to understand their behaviour more fully.

ACKNOWLEDGMENTS — The authors are grateful to Pat Hiddleston, Haydn Jones, Ian Selmes and Lorna Ogilvie who helped in arranging the collection of data, and to Alastair Pollitt and Gellisse Bagnall who carried out the computer analyses.

REFERENCES

BIGGS, J B (1978). Individual and group differences in study processes *Br. J. educ. Psychol.*, 48, 266-279.

BRISLIN, R. W (1983). Cross-cultural research in psychology. *Ann Rev Psychol*, 34, 363-400

ENTWISTLE, N J , THOMPSON, J. B , and WILSON, J. D. (1974). Motivation and study methods *Higher Educ* , 3, 379-396

ENTWISTLE, N J , and RAMSDEN, P (1983). *Understanding Student Learning* London:,Croom Helm.

ENTWISTLE, N J , and WILSON, J. D. (1977). *Degrees of Excellence: The Academic Achievement Game.* London. Hodder & Stoughton.

KOZÉKI, B. (1975) *Motivalas es Motivacio.* Budapest. Tankonyvkiado.

KOZÉKI, B (1980). A Motivalas es Motivacio Osszefuggeseinek. *Pedagogiai Pszichologiai Vizsgalata.* Budapest: Akademiai.

KOZÉKI, B (1984). Motives and motivational styles in education In ENTWISTLE, N J. (Ed), *New Directions in Educational Psychology. Learning and Teaching* Lewes: Falmer Press

KOZÉKI, B , and ENTWISTLE, N J (1983) Describing and utilizing motivational styles in education. *Br. J. educ. Stud*, XXXI, 184-197

KOZÉKI, B , and ENTWISTLE, N J. (1984) Identifying dimensions of school motivation in Britain and Hungary. *Br J educ Psychol*, 54, 306-319.

MARTON, F (1984) Approaches to Learning. In MARTON, F , HOUNSELL, D. J., and ENTWISTLE, N. J. (Eds.), *The Experience of Learning.* Edinburgh Scottish Academic Press.

MILLER, C M L , and PARLETT, M (1974). *Up to the Mark: A study of the Examination Game.* London. S R.H E.

NIE, N. H., *et al* (1975). *Statistical Package for the Social Sciences* New York: McGraw Hill

NIE, N. H., *et al* (1981). *Interactive Version of SPSS* New York: McGraw Hill

PASK, G. (1976) Styles and strategies of learning *Br. J. educ. Psychol*, 46, 128-148

SELMES, I. P (1985) Approaches to learning in secondary school their identification and facilitation Unpublished PhD thesis, University of Edinburgh.

WATKINS, D (1983) Assessing tertiary study processes. *Hum. Learn* , 2, 29-37

(*Manuscript received* 14*th January*, 1985)

Br J. educ. Psychol., 55, 138-147, 1985

IMPROVING PERSISTENCE THROUGH OBSERVATIONAL LEARNING AND ATTRIBUTION RETRAINING

BY M. L. CRASKE
(*University of Tasmania*)

SUMMARY. This study aimed to improve the persistence of learned helpless children through reattribution training accomplished via observational learning. In addition, the relationships of effort attributions and intelligence to persistence behaviour were examined. In the initial phase, 65 Grade 5 and 6 subjects were presented with unsolvable puzzle-completion problems to gauge persistence levels, they also completed the Intellectual Achievement Responsibility Scale (IARS), and Raven's Progressive Matrices. IARS scores which reflected neglect of the effort factor in academic outcome correlated highly with persistence; intelligence scores were not related to persistence. In the second phase, the 32 least persistent subjects were assigned to control or vicarious reattribution training groups. The treatment was associated with increased persistence in female, but not male, subjects The results are discussed in terms of exchange theory analysis and self-worth theory.

INTRODUCTION

THE understanding of persistence behaviour and other aspects of school achievement has been considerably advanced by the application of concepts drawn from attribution and learned helplessness theories. The general attributional approach to motivation developed by Weiner and his colleagues (e.g., Weiner, 1972, 1979) asserts that the causal beliefs people hold about their successes and failures have important consequences for their subsequent feelings, expectancies and behaviour. Learned helplessness is a specific instance of this general formula. It occurs when an individual believes he lacks control in obtaining desired outcomes, and when he believes the cause of his lack of control is an internal, stable and global factor. The consequences of such beliefs include diminished self-esteem, low success expectancy and deterioration in performance (Abramson *et al.*, 1978).

These concepts have been used to explain the deteriorated academic performance and reduced persistence levels of some students after repeated failure (e.g., Dweck and Reppucci, 1973; Dweck and Bush, 1976; Andrews and Debus, 1978; Diener and Dweck, 1980). For the student whose responses have consistently failed to produce the desired outcome (i.e., success), the situation appears uncontrollable. As a consequence, the student may become academically helpless. The causal ascription which has been shown to mediate such helplessness is the attribution of failure to lack of ability (Dweck and Reppucci, 1973; Diener and Dweck, 1978, 1980). Since ability is presumed to be relatively unchanging across time and relevant to a wide variety of situations and since its source lies within the individual, it can be described as an internal, global and stable factor. Children who respond to failure with improved performance and an escalation of persistence appear to be guarded against the effects of helplessness by attributing failure to lack of effort (Diener and Dweck, 1978). Effort is an unstable and specific factor. In addition, it is subject to the volitional control of the individual. Effort attributions therefore imply the possibility of better performances in the future and in other situations, and do not reflect on self-worth in the same way that ability attributions do (Butkowsky and Willows, 1980; Stipek and Hoffman, 1980).

Subsequent to recognition of the central role of attributions in achievement behaviour, a number of studies have attempted to improve persistence through attributional modification. Such an approach essentially aims to correct the maladaptive attributional pattern associated with helplessness, by encouraging children firstly to construe lack of effort as the cause of their failures, and, secondly, to recognise that success is more likely with effort. The efficacy of reattribution training has been demonstrated by Dweck (1975), Andrews and Debus (1978), Fowler and Peterson (1981) and Medway and Venino (1982).

The main purpose of this study is to assess the effectiveness of reattribution training when presented vicariously, rather than individually. There are a number of reasons for selecting observational learning as a potentially valuable means of producing attribution change, and so improving persistence. Firstly, behaviour change through vicarious methods usually occurs within a relatively short time period (Blackham and Silberman, 1975), therefore enhancing the cost-effectiveness of the procedure. Secondly, vicarious methods are generally more effective with observers who are low in self-esteem, highly dependent and incompetent (Blackham and Silberman, 1975). To some extent, these terms are descriptive of the helpless child who blames his own lack of ability for his failure and, though not incompetent generally, does not possess effective strategies to cope with increases in task difficulty. In addition, there is evidence that observational learning can influence those variables relevant to learned helplessness. Bandura (1977) has found modelling to be a successful means of changing self-efficacy expectations in phobic subjects, i.e., of inducing in the observer a belief that he is capable of performing a required behaviour, and therefore that his success in performing the behaviour is due to his own efforts. Brown and Inouye (1978), De Vellis *et al.* (1978), Breen *et al.* (1979), have all demonstrated performance decrements in undergraduate subjects who had not experienced failure. Zimmerman and Blotner (1979), and Zimmerman and Ringle (1981) noted improved persistence in first and second grade children who observed an adult model; the extent of improvement was dependent on the duration of the model's effort or performance, his/her success and his/her statements of confidence.

Ancillary to the central treatment aim of the study, two secondary predictions were made about the features associated with helplessness. In view of the large amount of research showing sex differences in susceptibility to learned helplessness (e.g., Crandall, 1969; Dweck and Reppucci, 1973; Le Unes *et al.*, 1980) and attributional tendencies (Dweck and Reppucci, 1973, Dweck and Bush, 1976; Dweck *et al.*, 1978; Nicholls, 1975, 1979), it was expected that females would demonstrate lower persistence levels than males, and that females would place less emphasis on the role of effort in explaining their failures.

It was also predicted that there would be an inverse relationship between susceptibility to helplessness and intellectual ability. Children whose intelligence lies in the retarded range are more likely to react to failure with reduced persistence than normals (Macmillan and Keogh, 1970, Weisz, 1979). This occurs because retarded children have a history of experiences in which their responses have largely failed to control outcomes, leading them to the conclusion that they are incapable of obtaining desired outcomes, and thus teaching them to be helpless. It seems likely that this paradigm will also apply, to a lesser degree, to children whose intelligence lies at the lower end of the normal range.

METHOD

Subjects

Thirty-seven male and 28 female pupils from two state primary schools near Hobart constituted the initial sample for the study. Mean ages were 10 years 11 months for the girls and 11 years 4 months for the boys. At school A, the subjects were Grade 5 and Grade 6 pupils from a composite Grade 5/6 classroom with a male teacher. At school B, the subjects were drawn from Grade 4/5 and Grade 5/6 composite units each of which shared a male and a female teacher. Only Grade 5 pupils from school

B participated in the study. The catchment areas of both schools are mixed in terms of socio-economic class but include a large percentage of working-class families.

Pretraining measures

Pretraining assessment consisted of measures of intellectual ability, effort attributions for failure and persistence. The persistence score was obtained by individual assessment; the remaining scores were derived from group administration of questionnaires completed in two sessions on consecutive days.

Intellectual ability. This was assessed using Raven's Progressive Matrices (PMS; Australian Council for Educational Research, 1958), a test of non-verbal intelligence. It is a timed test consisting of 60 incomplete designs, which the subject is required to complete by selecting missing sections from multiple choice arrays. Raw scores are converted to intelligence quotients. The Council for Educational Research (1958) has prepared Australian norms covering the range 10 to 18+ years, and reports high reliability scores and adequate validity.

Effort attributions. A subscale of the Intellectual Achievement Responsibility Scale (IARS; Crandall *et al.*, 1965) was used to obtain an index of the extent to which lack of effort was construed as a cause of failure. The IARS is a locus of control questionnaire designed to determine the degree to which a child believes that the intellectual failures and successes he/she encounters are a result of his/her own behaviour *versus* the behaviour of important others in his/her environment (such as teachers, parents and friends). It is comprised of 34 forced-choice items each depicting a positive or negative achievement situation and presenting two attributions to choose from: internal, in which responsibility for the outcome is assumed by the subject; or external, in which responsibility for the outcome is relegated to some property of the situation or other person. Crandall (1978) obtained a test-retest reliability of 0·76 for the IARS. Research with the scale indicates that helpless children take less personal responsibility for the outcomes of their behaviour and place less emphasis on the role of effort in determining success and failure than more persistent children (Dweck and Reppucci, 1973; Dweck, 1975; Crandall, 1978). A subscale of 10 items presents negative achievement situations and requires the subject to choose between his/her own lack of effort and an external factor as the causal attribution. Diener and Dweck (1978, 1980) found the subscale to distinguish between helpless and mastery-oriented subjects. In scoring the subscale one point is awarded for each effort attribution with a low score being associated with helplessness.

Persistence. An index of persistence, or susceptibility to learned helplessness, was obtained by assessing each subject's performance on a test of 15 unsolvable problems devised by the author. The top row of each problem contained a series of four patterns followed by a blank shape. The subject was required to fill in the blank shape and complete the sequence, by selecting from an array of four patterns given. The patterns were adapted from the Advanced Progressive Matrices, Set II (Raven, 1962). The instructions to each child were as follows:

> "We are going to do some puzzles. You will see on the first page of this book five shapes, the last one is blank. You have to find which one of the numbered shapes below fits into the blank space in the top row to follow on the pattern. If you look at the first one you can see that number 1 will fit into the top row to continue the pattern. When you decide which of the numbered shapes is the right one, write your answer on the sheet. Each time you write an answer a light will come on to tell you if you are right or wrong; a green light for right, and a red light for wrong. Don't change your answer if it is wrong, just go on to the next one. I would like you to do at least five puzzles; after that you can stop whenever you like. Close the book when you have finished all the puzzles you would like to do. You can start now."

The red light feedback was incorporated to enhance the salience of the failure result The Persistence score was obtained by summing the number of seconds spent on items 6 to 15. It was assumed that susceptibility to helplessness would be reflected in shorter times spent on each problem and a willingness to finish the test earlier, thus resulting in a lower Persistence score.

In order to provide some check on the success of the experimental manipulation, after completing the unsolvable problems, subjects were asked, "Do you think you made mistakes because I (the experimenter) was not telling you the truth about whether your answers were right or wrong?" None of the subjects answered affirmatively.

Finally, a set of eight easily solvable general knowledge questions was administered and the subject's success with these was stressed, to ensure that he did not leave the session with a feeling of failure.

The author was assisted by a female undergraduate psychology student, who administered the Persistence measure to approximately half of the subjects.

Selection of helpless subjects

Pretraining Persistence scores ranged between 0 and 291 seconds for boys, with a mean of 103·05 seconds; and between 0 and 167 seconds for girls with a mean of 78·75 seconds. Subjects who spent less than 80 seconds on the puzzles were selected as being least persistent and most susceptible to learned helplessness. Thirty-two of the 65 subjects fell into this category: 18 or 49 per cent of the boys, and 14 or 50 per cent of the girls. These 32 subjects were selected for participation in the second phase of the study. The sexes were randomly allocated to training and control groups, and paired according to sex, Persistence and PMS scores. (The latter were included in order to account for the possible influence of intelligence on the effectiveness of the training procedure.) Two boys were absent for the second phase, leaving training and control groups each consisting of seven girls and eight boys. The matched groups did not differ significantly on the Persistence ($t(14)=0{\cdot}45$, $P>0{\cdot}05$) or PMS measures ($t(14)=0{\cdot}29$, $P>0{\cdot}05$).

Training of helpless subjects

Training was directed at teaching the helpless subjects to reattribute failure to a lack of effort, rather than a lack of ability, through an observational learning procedure. To this end, an 8-minute training film was made in which a model was seen answering a set of 18 puzzles of the type used in the pretraining Persistence assessment. The model obtained success on these items at a rate of 33 per cent; the successes were arranged in a partial reinforcement schedule, with consecutive failures prior to success. There is evidence that this type of schedule is most effective in overcoming helplessness (Chapin and Dyck, 1976; Nation *et al.*, 1979; Nation and Cooney, 1980; Stein, 1980). The mechanism underlying the procedure's effectiveness is unknown, though possibly interlacing failure experiences with success enhances self-efficacy expectations.

In order to facilitate imitation, the model was rewarded for a correct answer by a green light and a token which could be exchanged for a selection of prizes at the end of the session. She/he was also told, "Good, you have chosen the right answer; that means you tried hard." Following a wrong answer, a red light was displayed and the model was told "That's the wrong answer; that means you should have tried harder."

Female experimental subjects observed a female model, while the male experimental subjects observed a male model. Both models were attractive 11-year-old children unknown to the subjects. Prior to viewing the films, the subjects were told:

"We are going to watch a film that will help you to do your work at school. You will see a girl (boy) who is very much like you — she (he) doesn't go to this school, but to a school near here, and she (he) is about the same age as you are. This girl (boy) is going to be doing some puzzles. Notice that when she (he) does not try very hard or spend much time on the question she (he) always gets the wrong answer. When she (he) does try hard and spends more time working out an answer she (he) always gets it right. Each time the answer is right she (he) earns a token to swap for a prize afterwards."

The subjects viewed the film for a second time, two days after the first showing. The control group saw two 8-minute excerpts from the television show "The Muppets" in place of the training film.

Post-training measure

Post-training assessment was undertaken on the day following the final film viewing. A second Persistence measure was obtained through repetition of the pre-training procedure using alternative, but similar puzzles. Again, the score was derived by summing the number of seconds the subject was willing to spend on items 6 to 15 in the face of consecutive failures, and again, none of the subjects indicated that they felt the experimenter was unreliable when giving failure feedback.

RESULTS

The predictions under investigation in the first phase of the study, prior to the selection of helpless subjects for intervention, concerned the relationships between persistence behaviour, attributional patterns and general intelligence. The existence of sex differences in persistence and attributional patterns was also the subject of examination. Table 1 presents the mean scores obtained for the variables assessed in the first phase.

TABLE 1

MEANS AND STANDARD DEVIATIONS OF SCORES ON PRETRAINING MEASURES

	Boys (N=37)		Girls (N=28)	
	M	SD	M	SD
PMS[a]	104 16	14 46	105 50	14 10
IARS[b]	6·87	1 83	5·32	2·02
Persistence[c]	103 05	66·88	78 75	46 00

[a] Scores represent intelligence quotients

[b] Maximum score = 10, high scores reflect attributions to lack of effort for failure

[c] Scores in seconds

The correlation between Persistence and IARS scores was found to be highly significant for both male ($r(35)=0{\cdot}69$, $P<0{\cdot}005$) and female subjects ($r(26)=0{\cdot}76$, $P<0{\cdot}005$). This supports the well-established association between willingness to persist in the face of consecutive failures, and an attributional disposition to construe failures as the result of lack of effort.

Intelligence levels, as assessed by the PMS, ranged between quotients of 82 and 134. About 87 per cent of the general population fall into this range (Australian Council for Educational Research, 1958). Those with high scores would be considered to have superior intelligence, and those with lower scores described as borderline (between dull-normal and mental retardation) The relationship between Persistence and PMS measures was found to be non-significant for male ($r(35)=0\cdot19$, $P>0\cdot05$) and female subjects ($r(26)=0\cdot22$, $P>0\cdot05$). Neither were significant associations found between IARS and PMS scores (Male: $r(35)=0\cdot02$, $P>0\cdot05$; Female: $r(26)=0\cdot12$, $P>0\cdot05$). It appears that intelligence, when lying essentially within the normal range, does not have a consistent influence on susceptibility to learned helplessness, in terms of behavioural persistence or attributional patterns.

Sex differences in susceptibility to learned helplessness were determined by comparing the Persistence and IARS scores of the 37 male and 28 female subjects. There was a trend for the girls to have lower Persistence scores than boys ($t(63)=1\cdot63$, $0\cdot10>P>0\cdot05$). A significant sex difference was noted in IARS scores ($t(63)=3\cdot17$, $P<0\cdot005$), with girls more likely to neglect the role of effort in causing their failures. Thus, while female subjects showed the attributional pattern indicative of greater susceptibility to learned helplessness, their behaviour was only marginally less persistent than that of male subjects.

In the second phase of the study, the central hypothesis concerned the effectiveness of vicariously presented reattribution training in improving persistence. Table 2 shows the mean persistence scores of the 32 least persistent subjects both prior to and following the intervention.

TABLE 2

MEANS AND STANDARD DEVIATIONS OF PERSISTENCE SCORES OBTAINED AT PRE- AND POST-TRAINING ASSESSMENT, FOR TRAINING AND CONTROL GROUPS (IN SECONDS)

		Training		Control	
		Pre	Post	Pre	Post
Boys	M	48 50	63·50	44 25	48·88
	SD	27 20	66·14	27·60	65 17
Girls	M	40·86	132·71	40 00	16 86
	SD	28 28	36·44	24·05	16·22

The pretraining Persistence scores of these subjects and those of the more persistent subjects who did not take part in the second phase differed significantly. (Males: $t(35)=7\cdot64$, $P<0\cdot0005$; Females: ($t(26)=7\cdot68$, $P<0\cdot005$).

Treatment effects were evaluated at post-training using a covariance analysis with pretraining Persistence as the covariate. The analysis showed a significant main effect for Experimental Condition ($F(1,25)=11\cdot63$, $P<0\cdot01$) and also a significant sex by Experimental Condition interaction ($F(1,25)=7\ 96$, $P<0\cdot01$). No significant main effect was found for sex ($F(1,25)=1\cdot64$, $P>0\cdot05$). Post-hoc analyses using t-tests to compare adjusted group means showed that the difference between treatment conditions was significant for female subjects ($t(12)=4\cdot56$, $P<0\cdot001$) but not for male subjects ($t(14)=0\cdot43$, $P>0\cdot05$). The vicarious reattribution training was therefore effective in improving Persistence in females only.

DISCUSSION

The major tenet of the attributional model of achievement motivation is that causal ascriptions influence and perhaps even determine subsequent achievement behaviours. The correlation between level of persistence and responses to the IAR subscale tends to support this. A high degree of persistence was associated with attributions for failure to lack of effort. In addition to the confirmation of past research (e.g., Dweck and Reppucci, 1973; Andrews and Debus, 1978; Diener and Dweck, 1978) this finding adds strength to the rationale of improving persistence through a reattribution approach. That observational learning is a successful means of enhancing persistence received only partial support in this study. Observation of a model who was rewarded for attributing failure to lack of effort was paralleled by a significant increase in persistence in helpless female subjects, but failed to produce improvements in helpless male subjects.

The differential effectiveness of observational learning is consistent with proposals that males and females differ in the extent to which they draw on environmental factors in performance evaluation. The literature on field dependence, for instance, shows that the perceptions of males are more independent of contextual cues than the perceptions of females (Witkin and Goodenough, 1977). Similarly, Rothbaum's (1981) work on control and responsibility has shown that males are more concerned with what they are able to do irrespective of others, whereas females are more concerned with what other persons can do. To account for these sex differences, it has been proposed (Crandall, 1963; Veroff, 1969; Crandall *et al.*, 1964) that during childhood boys develop internal standards of evaluation which allow them to become relatively independent of feedback from others, while girls remain dependent on external evaluation to assess the quality of their performance.

Assuming that boys and girls do differ in this way, then it is possible to explain the sex difference in treatment effectiveness using an extension of the exchange theory analysis of achievement behaviour (Wyer and Bednar, 1967). According to this analysis the observer of a model receiving rewards for his behaviour compares himself with the model, and acquires expectations for his own reward or success based on the degree of the model's success. When the observer subsequently participates without success, his degree of tolerance for failure (i.e., his willingness to persist) is related to the expectations he has developed. A low success expectancy will be acquired if the observer views a model with a low success rate, such as the 33 per cent rate of correct answers employed in this study. This low success expectancy will result in a high frustration tolerance and high level of persistence when the observer is faced with consecutive failures himself. Given that female observers have a greater need than male observers for external evaluation, their expectancies and subsequent behaviours will be modified to a larger extent by the social comparison process. Thus, girls may demonstrate higher levels of persistence than boys after observing a model with a low rate of success.

An alternative and perhaps more straightforward explanation of the sex difference found can be derived from Beery's (1975) self-worth theory. Self-worth theory asserts that achievement behaviour is largely determined by student attempts to maintain a self-concept of high ability. From this perspective, a combination of high effort expenditure and failure is potentially threatening to the student since it increases suspicions of low ability. Just such a combination was presented to subjects in post-training assessment; they were asked to tackle unsolvable problems after being trained to associate effort with achievement and lack of effort with failure. Research into self-worth motivation suggests that males are more accomplished at self-aggrandisement and more concerned with maintaining an appearance of competence than females (Covington and Omelich, 1979; Covington *et al.*, 1980). This greater desire for the protection of self-worth may explain why the boys were less willing to expend effort at post-test than the girls, thus achieving lower Persistence scores.

If boys are prevented from expending effort when risking failure because of self-worth motivation, the efficacy of effort attribution training with helpless boys is likely to be severely restricted. However, establishing the efficacy of vicariously presented reattribution training still represents an important contribution to the treatment of female helplessness. The advantages of vicarious procedures lie in their cost-effectiveness: group rather than individual training is possible, and behaviour change usually occurs after relatively short periods of observational learning. In this study, only two sessions of approximately 8 minutes' length produced improvements in persistence. The procedure therefore becomes feasible as a remedial device within the classroom, for female students. It is essential that future studies investigate the generalisability and durability of the persistence improvement, and its effects on academic performance.

One final issue is important in considering the behaviour and cognitions of the least persistent subjects. The possibility exists that some children may have opted out, or become less persistent, when presented with unsolvable problems because they realised they were being duped by the experimenter. If this were the case, then the experimental conditions would have little validity in reflecting real life helplessness situations, in which solutions are possible. In the present sample, none of the subjects indicated they felt they were being misled by the experimenter when they were presented with consecutive red lights in the Persistence tasks. This implies that they had faith in the failure feedback, but does not reflect on their faith in the problems themselves. Some subjects may have had suspicions about their unsolvable nature. The experimental design does not resolve this issue.

In view of the prevalence of helplessness in retarded children (Macmillan and Keogh, 1970; Weisz, 1979), the finding that low-normal intelligence does not predispose the individual to helplessness was surprising. Presumably, the performance history of the low-normal IQ subjects in this study was not dominated by failure to the extent that they learned to view themselves as incapable of achieving success. An issue yet to be investigated is the relationship between superior intelligence and persistence. It may be that particularly bright children are as vulnerable to helplessness as retardates since their performance histories would involve little contact with failure, and therefore few opportunities to develop strategies to cope with failure. Experimentally, the 'success only' reinforcement schedule has been ineffective or less effective than partial reinforcement in enhancing persistence (Nation *et al.*, 1979).

The finding most discrepant with previous research was the similarity of male and female persistence levels. There has been general agreement in the literature that girls are less persistent and more helpless than boys in the 10 to 12 year age group (e.g., Crandall, 1969; Dweck and Reppucci, 1973). The present results indicated only a marginal trend for females to be less persistent, despite the fact that they displayed the attributional pattern associated with helplessness to a greater extent. The role of teacher behaviour in this context is worthy of some scrutiny. Both Crandall (1978) and Dweck *et al.* (1978) argue that sex differences in persistence and attributions are derived largely from the type of feedback given by primary school teachers. Classroom observations led them to conclude that since boys receive a greater proportion and a more indiscriminant application of negative feedback they are discouraged from viewing failure as indicative of low ability. However, such differential reinforcement contingencies have not been found to operate in all classrooms (Heller and Parsons, 1981; Meece *et al.*, 1982) and it appears that they are more likely to operate in the classrooms of female teachers (Sikes, 1971). Furthermore, Dweck and Bush (1976) reported that failure feedback from a male adult agent led to less helplessness than similar feedback from a female adult.

In this study, male teachers had either total or 50 per cent responsibility for the classes from which the subjects were drawn. It is possible therefore that the contingencies assumed to promote sex differences in helplessness were not in existence, or were diminished in their effectiveness. Since female teachers predominate in primary schools, and since generalised attributions for failure are presumed to mediate sex differences in long-term academic achievement, the issue of differential feedback styles clearly warrants further investigation.

ACKNOWLEDGMENTS. — The author is thankful for the assistance of Gabrielle Craske, and comments by Iain Montgomery and Gemma O'Callaghan. Requests for reprints should be sent to M. L. Craske, Department of Psychology, University of Tasmania, Hobart, Tasmania 7000, Australia.

REFERENCES

ABRAMSON, L Y, SELIGMAN, E. P., and TEASDALE, J D. (1978). Learned helplessness in humans. critique and reformulation *J. abnorm Psychol.*, 87, 49-74.

ANDREWS, G. R, and DEBUS, R. L. (1978). Persistence and the causal perception of failure· modifying cognitive attributions *J. educ. Psychol*, 70, 154-166.

Australian Council for Educational Research (1958). *Manual for Standard Progressive Matrices.*

BANDURA, A. (1977) Self efficacy toward a unifying theory of behavioural change. *Psychol. Rev.*, 84, 191-215.

BAR-TAL, D., RAVIV, A., RAVIV, A., and BAR-TAL, Y (1982). Consistency of pupils' attributions regarding success and failure. *J. educ. Psychol.*, 74, 104-110.

BEERY, R (1975). Fear of failure in the student experience. *Personnel and Guidance J.*, 54, 190-203.

BLACKHAM, G. J., and SILBERMAN, A. (1975) *Modification of Child and Adolescent Behaviour.* Belmont· Wadsworth

BREEN, L. J., VULCANO, B, and DYCK, D. (1979) Observational learning and sex roles in learned helplessness. *Psychol Rep*, 44, 135-144

BROWN, I, and INOUYE, D. K (1978). Learned helplessness through modelling: the role of perceived similarity in competence. *J. Person. soc. Psychol.*, 36, 900-908.

BUTKOWSKY, I. S., and WILLOWS, D. M (1980). Cognitive-motivational characteristics of children varying in reading ability: evidence of learned helplessness in poor readers *J. educ. Psychol.*, 72, 408-422.

CHAPIN, M., and DYCK, D. G. (1976). Persistence in children's reading behaviour as a function of N length and attribution retraining *J abnorm Psychol.*, 85, 511-515

COVINGTON, M. V, and OMELICH, C. L. (1979). It's best to be able and virtuous too: student and teacher evaluate responses to successful effort. *J educ Psychol*, 71, 688-700.

COVINGTON, M. V, SPRATT, M F, and OMELICH, C. L. (1980). Is effort enough, or does diligence count too? Student and teacher reactions to effort stability in failure. *J educ Psychol*, 72, 717-729

CRANDALL, V. C. (1969) Sex differences in expectancy of intellectual and academic reinforcement In SMITH, C P. (Ed), *Achievement-related Motives in Children.* New York: Russell Sage Foundation.

CRANDALL, V. C. (1978). New developments with the Intellectual Achievement Responsibility Scale. Symposium· Goal-Specific Locus of Control Scales, American Psychological Association Convention, Toronto, 1978.

CRANDALL, V C, KATKOVSKY, W, and CRANDALL, V J. (1965) Children's beliefs in their own control of reinforcements in intellectual-academic situations. *Child Dev.*, 36, 91-109.

CRANDALL, V. J (1963). Achievement In STEVENSON, H. W. (Ed.), *The 62nd Yearbook of the National Society for the Study of Education (Pt. 1) Child Psychology.* Chicago. National Society for the Study of Education.

CRANDALL, V. J, DEWEY, R., KATKOVSKY, W., and PRESTON, A. (1964). Parents' attitudes and behaviours and grade school children's academic achievements. *J genet Psychol.*, 104, 53-66.

DE VELLIS, R. F., DE VELLIS, B M., and MCCAULEY, C (1978) Vicarious acquisitions of learned helplessness. *J. Person soc Psychol.*, 36, 894-899.

DIENER, C., and DWECK, C S (1978). An analysis of learned helplessness continuous changes in performance, strategy, and achievement cognitions following failure *J Person, soc. Psychol.*, 36, 451-462.

DIENER, C, and DWECK, C. S (1980) An analysis of learned helplessness II. The processing of success. *J Person soc. Psychol.*, 39, 940-952

DWECK, C. S (1975). The role of expectations and attributions in the alleviation of learned helplessness *J Person. soc Psychol.*, 31, 674-685

DWECK, C. S. and BUSH, E S. (1976). Sex differences in learned helplessness. I Differential debilitation with peer and adult evaluators. *Dev Psychol*, 12, 147-156

DWECK, C. S., DAVIDSON, W, NELSON, S., and ENNA, B. (1978). Sex differences in learned helplessness. II The contingencies of evaluative feedback in the classroom. III. An experimental analysis. *Dev. Psychol*, 14, 268-276.

DWECK, C S., GOETZ, E. E., and STRAUSS, N L (1980). Sex differences in learned helplessness· IV. An experimental and naturalistic study of failure generalization and its mediators. *J. Person. soc. Psychol.*, 38, 441-452.

DWECK, C. S, and LICHT, B. C (1980). Learned helplessness and intellectual achievement. In GARBER, J. and SELIGMAN, M. E. P (Eds), *Human Helplessness — Theory and Applications* New York· Academic Press

Dweck, C S., and Reppucci, D. (1973) Learned helplessness and reinforcement responsibility in children. *J Person soc. Psychol.*, 25, 109-116.

Fowler, J. W., and Peterson, P L (1981). Increasing reading persistence and altering attributional style of learned helpless children. *J educ Psychol.*, 73, 251-260.

Heller, K. A., and Parsons, J E. (1981). Sex differences in teachers' evaluative feedback and students' expectancies for success in mathematics *Child Dev.*, 52, 1015-1019

Le Unes, A. D , Nation, J R., and Turley, N. M (1980). Male–female performance in learned helplessness. *J. Psychol.*, 104, 255-258

MacMillan, D L , and Keogh, B. K. (1970). Normal and retarded children's expectancy for failure. *Dev. Psychol*, 4, 343-348.

Medway, F J., and Venino, G R. (1982). The effects of effort feedback and performance patterns on children's attribution and task persistence *Contemp. educ Psychol*, 7, 26-34

Meece, J L , Parsons, J. E., Kaczala, C M , Goff, S. G , and Futterman, R. (1982). Sex differences in math achievement· toward a model of academic choice. *Psychol Bull.*, 81, 324-348

Nation, J L , and Cooney, J. B. (1980). The change and maintenance effectiveness of persistence training regarding the treatment of laboratory-induced and naturally occurring depression *Bull Psychonomic Soc*, 16, 121-124

Nation, J L., Cooney, J B , and Gartrell, K. E. (1979) Durability and generalizability of persistence training. *J. abnorm Psychol.*, 88, 121-136.

Nicholls, J. G (1975). Causal attributions and other achievement-related cognitions effects of task outcome, attainment value, and sex *J Person soc. Psychol*, 31, 379-389

Nicholls, J. G (1979) Development of perception of own attainment and causal attributions for success and failure in reading *J. educ. Psychol.*, 71, 94-99

Raven, J (1962) *Advanced Progressive Matrices.* London Lewis

Rothbaum, F. (1981). The effect of sufficiency and necessity on perceptions of control and responsibility *J Personality*, 49, 85-100.

Sikes, J (1971) Differential behaviour of male and female teacher with male and female students Unpublished doctoral dissertation, University of Texas at Austin

Stein, N (1980) Inoculation against learned helplessness. *Psychol Rep.*, 47, 1143-1151

Stipek, D. J., and Hoffman, J M. (1980). Children's achievement-related expectancies as a function of academic performance histories and sex. *J educ Psychol.*, 72, 861-865

Veroff, J. (1969) Social comparison and the development of achievement motivation In Smith, C P (Ed), *Achievement-related Motives in Children* New York. Russell Sage

Weiner, B. (1972) *Theories of Motivation From Mechanism to Cognition.* Chicago: Rand McNally.

Weiner, B (1979). A theory of motivation for some classroom experiences *J. educ. Psychol*, 71, 3-25

Weisz, J R. (1979). Perceived control and learned helplessness among mentally retarded and nonretarded children: a developmental analysis. *Devel. Psychol.*, 15, 311-319.

Witkin, H. A., and Goodenough, D. R. (1977) Field dependence and interpersonal behaviour. *Psychol Bull*, 84, 661-689

Wyer, R. S., and Bednar, R. (1967). Some determinants of perseverance in achievement-related activity *J exp soc. Psychol.*, 3, 255-265

Zimmerman, B J , and Blotner, R. (1979) Effects of model persistence and success on children's problem-solving *J. educ Psychol*, 71, 508-513.

Zimmerman, B J., and Ringle, J (1981) Effects of model persistence and statements of confidence on children's self-efficacy and problem-solving *J. educ. Psychol*, 73, 485-493.

(*Manuscript received* 12*th September*, 1984)

Br. J. educ. Psychol., 55, 148-155, 1985

CONCEPTUALISING RESEARCH ON EFFECTIVE TEACHING

By CHRIS KYRIACOU
(Department of Education, University of York)

SUMMARY. A framework is presented for considering effective teaching which focuses on 'classroom process'. The framework distinguishes three complementary levels of analysis. First, a surface level of analysis which focuses on the notions of 'active learning time' and 'quality of instruction'. Second, a psychological level of analysis which focuses on psychological concepts, principles and processes which attempt to explain *how* process may influence outcomes. Third, a pedagogical level of analysis which focuses on teaching skills and on a three-way interaction between teacher perceptions and strategies, pupil perceptions and strategies and characteristics of the learning experience. It is argued that the distinction between these three levels of analysis may aid clarification in conceptualising research on effective teaching.

INTRODUCTION

IN this paper an attempt is made to provide a framework within which the notion of effective teaching can be considered. The starting point for this consideration is 'the lesson' — the unit of classroom teaching. At its crudest level, the central question being addressed here is 'What is the most effective lesson that can be given by a teacher for a specified context and with a specified set of outcomes desired?" As such, the main focus of attention here will be on teacher and pupil behaviour during a lesson.

The specified *context* of a lesson here refers to all those variables which may have an influence on teacher and pupil behaviour during the lesson, such as teacher characteristics (e.g., experience, knowledge), pupil characteristics (e.g., age, ability), class characteristics (e.g., range of ability, size), subject characteristics (e.g., topic, level of difficulty), school characteristics (e.g., ethos, size), community characteristics (e.g., urban/rural, affluence) and characteristics of the occasion (e.g., weather, time of day).

The specified set of desired *outcomes* of a lesson here refers to the outcomes for pupils of a lesson which can be regarded as the criteria for judging its effectiveness. The most important of these appear to be increased knowledge and skills, increased interest in the subject or topic, increased intellectual motivation, increased academic self-confidence and self-esteem, increased development of pupil autonomy, and increased social development. For each lesson, and for a course as a whole, the teacher will implicitly or explicitly have in mind a ranking of the relative importance of these outcomes. In addition, this ranking may well change significantly from lesson to lesson, and may also vary somewhat at different times during the lesson and for different subgroups of pupils within the class.

Process variables here refer to teacher and pupil behaviour in the lesson, and include teacher's clarity, enthusiasm, use of questions, use of praise and criticism, use of non-verbal communication, management strategies, "withitness", disciplinary techniques, classroom climate, organisation of the lesson, suitability of learning tasks, type of feedback pupils receive, pupil involvement in the lesson, pupil-initiated interaction with the teacher, and pupils' strategies for learning. This list is by no means exhaustive. Many such variables have been operationalised in terms of discrete observable behaviours which can be monitored and a number of attempts have been made to summarise their relationship with different outcomes (e.g., Rosenshine, 1971; Anderson *et al.*, 1980; Good and Brophy, 1980; Ryan and Butler, 1982); Kyriacou, 1983).

Almost all research on effective teaching reported over the last few decades has been considered within this basic framework of Context-Process-Outcomes, and, although it has been elaborated in a number of ways (see Bennett, 1976, 1978; Winne and Marx, 1977; Rutter *et al.*, 1979; Centra and Potter, 1980; Kyriacou and Newson, 1982; Haertel *et al.*, 1983), there nevertheless appears to be little to gain from attempting to develop this basic framework further. These overviews have by no means been successful in identifying an unequivocal set of process variables which have a consistent influence on a wide range of outcomes across a wide range of contexts. However, some degree of reliability of findings has emerged once the specified outcome and context has been taken into account (see e.g., Eggleston *et al.*, 1976). Thus, in response to a question such as "Are enthusiastic teachers more effective?", the guarded answer that has emerged is probably that it depends on the context and the outcome being considered, and that one needs to bear in mind that the relationship may not be a simple linear one, that it may be affected by the presence of other process variables, and that it may have differing effects on different subgroups of pupils. The problems facing research within the framework of Context-Process-Outcomes have been much discussed (e.g., Kyriacou and Newson, 1982).

During the last two decades increasing attention has been paid to the notion of 'process', and it is here that a number of interesting developments have taken place. These developments have largely stemmed from a recognition that early research on process: firstly, had not yielded unequivocal and consistent findings; secondly, had not adequately taken account of the influence of teacher and pupil perceptions of the learning experience and their managerial, coping and learning strategies on outcomes; and, thirdly, had not adequately attempted to explain *how* process variables have an influence on outcomes.

Contemporary thinking about effective teaching focusing on process thus seems to be best described with reference to three ways of conceptualising effective teaching; these are illustrated in Figure 1. These three conceptualisations should be seen as complementary and in essence are consistent with each other. In effect, they are three ways of looking at and 'unpacking' the same phenomena.

THREE LEVELS OF ANALYSIS

A surface level of analysis

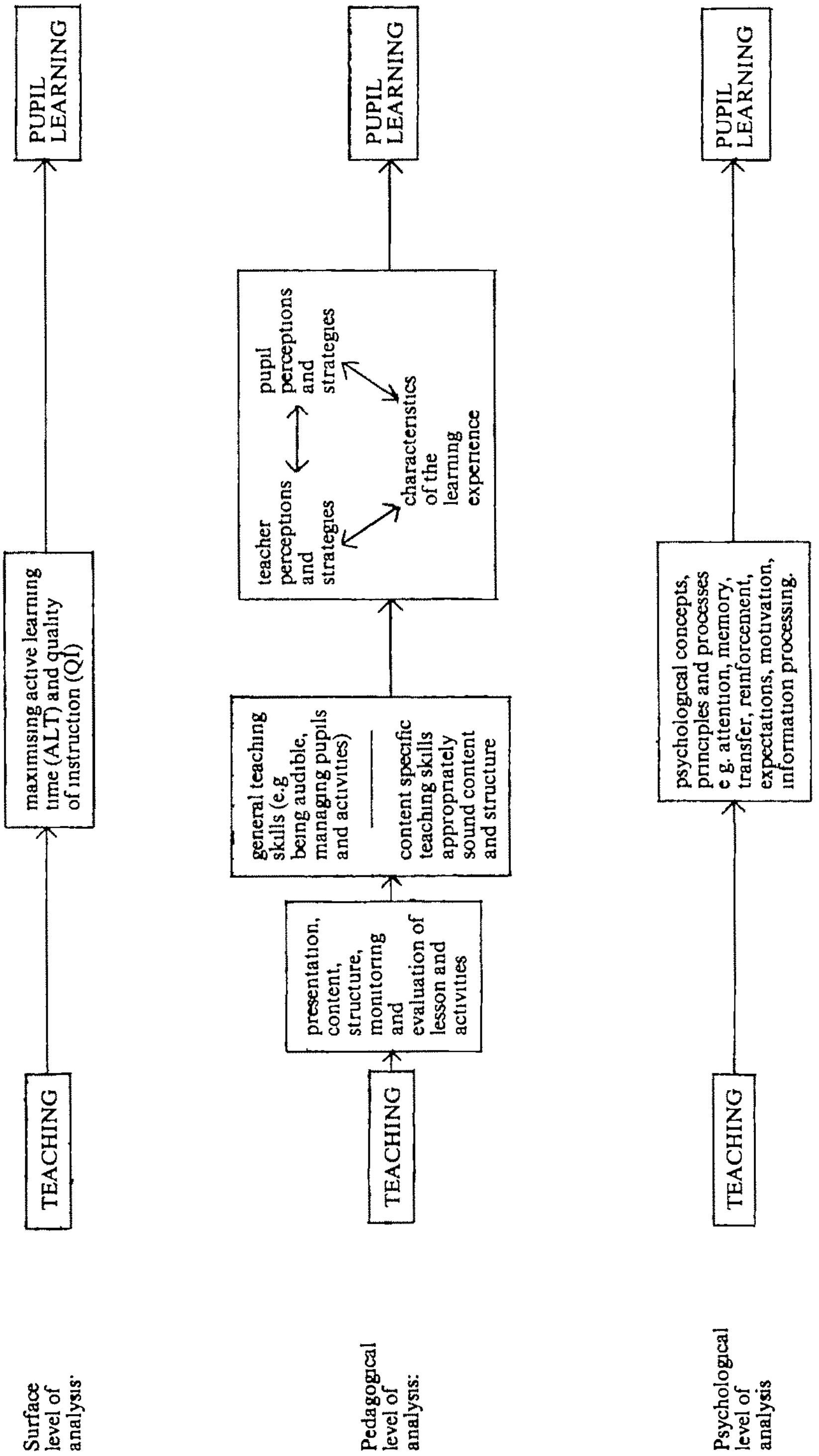
Recent research and theorising has distinguished two complementary foci of attention in considering process variables (see Peterson and Walberg, 1979; Centra and Potter, 1980; Bennett and Desforges, 1983; Anderson, 1984). The first is *Active Learning Time* (ALT), and refers to the amount of time spent by a pupil during a lesson actively engaged in outcome related behaviour (either by interacting with the teacher or by undertaking a task) (see Fitz-Gibbon and Clark, 1982). The second is *Quality of Instruction* (QI), and refers to the extent to which the intereaction with the teacher and the tasks undertaken facilitate the development of the desired outcomes (see Corno, 1979). It is argued within this level of analysis that effective teaching (and hence the influence of process variables on outcomes) is based on maximising ALT and QI during the lesson.

Although some process variables may appear to influence outcomes primarily through either ALT or QI, it is important to note that the vast majority of process variables contribute to both ALT and QI. Indeed, this point is implicit in much of the research on classroom management (see Kounin, 1970; Robertson, 1981; Freiberg, 1983). Thus, although it is argued that increasing ALT (for a given level of QI) or increasing QI (for a given level of ALT) will increase effectiveness, it is important to take into account both ALT and QI in considering effectiveness.

Of particular interest in developments here has been the gradual modification of

FIGURE 1

THREE WAYS OF 'UNPACKING' PROCESS


Surface level of analysis:
TEACHING
maximising active learning time (ALT) and quality of instruction (QI)
PUPIL LEARNING
Pedagogical level of analysis:
TEACHING
presentation, content, structure, monitoring and evaluation of lesson and activities
general teaching skills (e.g. being audible, managing pupils and activities)
content specific teaching skills appropriately sound content and structure
teacher perceptions and strategies
pupil perceptions and strategies
characteristics of the learning experience
PUPIL LEARNING
Psychological level of analysis
TEACHING
psychological concepts, principles and processes e.g. attention, memory, transfer, reinforcement, expectations, motivation, information processing.
PUPIL LEARNING

these two concepts such that they now attempt to encompass in some way the pupil's experience of learning in terms of being cognitively and affectively engaged in the fullest sense (Anderson, 1984; Bennett *et al.*, 1984).

A psychological level of analysis

In order to fully understand the influence of process variables on outcomes, this influence needs to be described in terms of various psychological concepts, principles and processes, such as attention, memory, transfer, reinforcement, mental set, expectations, attitudes, self-concept, cueing, feedback, concept discrimination, motivation, identification, ability and information processing. (This list is by no means exhaustive, and emphasis will differ somewhat depending on the particular psychological theories and models adopted: see Good and Brophy, 1980; Tomlinson, 1981; Haertel *et al.*, 1983; Howe, 1984).

In essence, this level of analysis seeks to make explicit the relationship between process variables and the psychological conditions necessary for the learning to occur; these psychological conditions furthermore need to be explicable at the level of each individual pupil in terms of that pupil's psychological state and in accordance with the nature of learning itself.

In considering the pupil's psychological state and the nature of learning, a number of models of 'instructional psychology' have been developed (see Haertel *et al.*, 1983). The present author sees three aspects as crucial to pupil learning in the context of effective teaching: (1) the pupil must be *attending* to the learning experience, (2) the pupil must be *receptive* to the learning experience (in the sense of being motivated or having a willingness to learn or respond to the experience), and (3) the learning experience must be *appropriate* for the desired learning to take place (taking particular account of the pupil's initial knowledge and understanding).

These three aspects of learning can be elaborated in a number of ways depending on the model of instructional psychology adopted, and indeed other authors may highlight other crucial aspects. However, the central point being made here is that such models, utilising psychological concepts, principles and processes, attempt to link process variables to the psychological conditions necessary for learning to take place, and as such attempt to offer an explanation of *how* process variables influence pupil learning.

Within such models a useful distinction has emerged in recent theorising between general conditions necessary for learning (common to all individuals) and conditions which are discussed primarily in terms of individual differences (e.g., pupil's degree of control over the learning experience); these latter conditions appear to act as 'catalysts' in that they attempt to explain individual differences between pupils that moderate the extent to which the general conditions for learning are achieved (rather than constituting a necessary condition in themselves) (see e.g., Gammage, 1982). Such theorising would suggest that whilst some learning will take place for all individuals if the general conditions for learning are met for those individuals, the desired learning will be more effectively fostered for each individual if the psychological processes elicited by a learning experience are tailored to that individual pupil's psychology (cf. preferred learning style) and to the particular type of outcome desired.

A pedagogical level of analysis

The pedagogical level of analysis has emerged from analyses of effective teaching which have concerned themselves directly with the needs of teacher education. It has largely evolved from attempts to explore the characteristics of effective teaching within a framework that sees teaching as a managerial activity, and has thereby sought to identify the central 'teaching skills' involved. Such an approach has tended to rely on 'common sense' acceptability to student and established teachers, depending more on its usefulness within teacher training and in-service education for its validity that on its explicit derivation from either research findings on effective teaching or models of instructional psychology.

This approach, based on the notion of teaching as a managerial activity, acquired a major influence on researchers (in addition to teacher educators) following the work of Kounin (1970). Kounin appeared to have identified a number of managerial techniques employed by 'effective' teachers (e.g., 'withitness', 'overlapping', 'smoothness', 'momentum') which seemed to have major implications for theorising at the surface level and psychological level of analyses. What was of particular importance concerning these findings was that they focused attention on qualities which were subjectively richer in meaning and execution than those typically employed as process variables.

Kounin's research was essentially conducted within the framework which sought to identify teaching skills. More recent writings, however, have moved significantly beyond the identification of teaching skills. What now seems to be the focus of attention is the attempt to explore the complex relationship between three aspects of process: the teacher's perceptions, decision-making, managerial and coping strategies; the pupil's perceptions, coping and learning strategies; and characteristics of the learning tasks, activities and experience (see Figure 1). These three aspects of process are then related to pupil learning (see Doyle, 1983, 1984; Bennett *et al.*, 1984; Calderhead, 1984; Doyle and Rutherford, 1984; Wragg, 1984). It is this three-way interaction between teacher perceptions and strategies, pupil perceptions and strategies, and the learning experience, which characterises a profound and important shift in current conceptualisations about process.

The pedagogical level of analysis thus seems to involve two main foci of attention as mediators between teaching and pupil learning. The first focus of attention is the attempt to unpack teaching into a number of managerial skills concerning the main tasks of teaching: presentation, content, structure, monitoring and evaluation of the lesson and its activities (see e.g., Waterhouse, 1983). A number of researchers and tutors involved in teacher education have attempted to list certain 'does' and 'don'ts' regarding these tasks. In considering the managerial skills discussed here, a useful distinction can be made between *general teaching skills* (such as being audible, and how to manage pupils and activities) and *content specific teaching skills* (which are concerned with the appropriateness of the content, method and structure adopted for the particular desired learning in hand). These two sets of teaching skills are complementary. Whilst some skills (e.g., audibility) seem to be clear examples of the former and other skills (e.g., ability to structure a lesson on a particular topic such as conservation of momentum into intellectually sound constituent parts) seem to be clear examples of the latter, a number of skills may integrate the two sets (e.g., teacher's clarity of exposition may involve both audibility and intellectually sound structure). A number of attempts have been made to identify the extent of individual differences between teachers on 'key teaching skills or qualities'. For example, Kyriacou and McKelvey (1984) employed a set of eight rating scales (each anchored by a description of behaviours at the end of the scales) in order to compare lessons in terms of eight 'key characteristics of teaching': 'preparedness', 'pace and flow', 'transitions', 'cognitive matching', 'clarity', 'business-like', 'withitness' and 'encouragingness'. What was of particular interest to the present paper was that this study indicated that whilst teachers may display similar levels of certain skills that does not adequately tap the diversity of practice utilising these skills. Indeed, it is this very point (that differences between lessons in terms of the types of skills displayed by teachers is not the full story) that has led to the shift of attention to the second main focus of attention.

The second focus of attention is the attempt to explore the three-way interaction between teacher perceptions and strategies, pupil perceptions and strategies, and the learning experience itself noted above. Interestingly, much of the recent research here (e.g., Doyle, 1983, 1984; Bennett *et al.*, 1984; Calderhead, 1984) can be seen to be a 'psychological' version of the sort of research conducted by sociologists who have employed an ethnographic approach to the exploration of classroom process in order to identify the meaning for the participants (pupils and teachers) of what was happening in the classroom (see Woods, 1980a, 1980b). The research of Bennett *et al.* and Doyle has highlighted how pupils actively attempt to make sense of what is required of them in a lesson, and it is in the final analysis how pupils react to the teaching and the learning experience that forms the link between the lesson and their learning. Observation of pupils and questioning of them has indicated that pupils may often interpret the purpose of a lesson and the learning tasks in ways clearly not intended by the teacher; in part this stems from assumptions by a teacher of how the lesson will be experienced which are simply wrong (particularly by assuming levels of motivation, understanding and competence in study skills which are too high, and by underestimating the extent to which pupils actively construct and employ algorithms, viz. simple rules and routines, which enable them to produce what is apparently demanded of them in a way that may be incorrect or undesirable).

The implications for effective teaching of this second focus of attention is that the teacher needs to employ a method of monitoring and evaluating pupil learning which can highlight pupil perception, coping and learning strategies and thereby mitigate some of the problems which accrue from certain assumptions about how a lesson will be experienced by pupils. Such an approach needs to take account of ways in which teachers can 'control' or utilise in some sense pupils' perceptions to advantage. Interestingly, discussions of pupil perception, coping and learning strategies seem to be best described in terms of illustrative cases and anecdotes, and it may be difficult to identify a particular type of teaching skill which could clearly relate to ways in which teachers could control or utilise such perceptions, although the notion of social sensitivity (the ability to see things from another's perspective) may be involved here.

CONCLUSIONS

The framework developed here, in identifying three complementary levels of analysis (illustrated in Figure 1), enables one to locate the rationale and findings of different studies within one overall framework and thereby to indicate how the assumptions and findings of such studies may relate to each other. A major problem facing research on effective teaching is the complex inter-relationship and confounding between different variables; it is hoped that the three ways of conceptualising the analysis of effective teaching developed here may usefully aid clarification by indicating how certain variables are primarily located in one of the three levels of analysis.

It is likely that the notion of 'teaching skills or qualities' will continue to receive much attention in research on effective teaching, particularly in the United Kingdom and in the United States where a number of teacher appraisal and teacher competence schemes are in progress. The present framework indicates that teaching skills and qualities can be usefully located within a pedagogical level of analysis; its effects on outcomes can be considered however in three ways. First, each skill or quality can be explored in terms of its contribution to maximising ALT and/or QI (surface level analysis); second, it can be explored in terms of psychological concepts, principles and processes with particular reference to how each skill or quality fosters and maintains the appropriate conditions for pupil learning to occur (psychological analysis); third, it can be explored in relation to teacher perceptions and strategies, pupil perceptions and strategies and characteristics of the learning experience itself (pedagogical analysis). Although the three levels of analysis are illustrated in Figure 1 as operating in parallel, their complementary nature needs to be borne in mind, and it can be seen that a number of important points about effective teaching depend on making links between these three levels.

As has been noted earlier, a significant shift in attention at the pedagogical level of analysis has taken place in recent years towards looking at the interaction between teacher perceptions and strategies, pupil perceptions and strategies and the learning experience itself. Within teacher education in particular, such research may enable student teachers to be made explicitly aware of those techniques or skills they can employ to identify and monitor their own teaching, the perceptions and strategies employed by pupils, and characteristics of the learning experience as experienced by the pupils.

In addition, there is a need to look at the notion of 'teaching styles' in a flexible way that takes account of the diversity of packages of teaching skills and strategies which may be equally effective and which take account of the holistic quality of teaching and how each style may be experienced by pupils; such an approach would place emphasis on general qualities concerning the teaching as experienced by pupils rather than attempt to identify a large number of discrete teaching behaviours which were then grouped together to form the basis for describing different styles.

Finally, much of the research conducted on effective teaching implicitly sees the teacher as involved in whole class instruction and as a transmitter of knowledge with the pupil in a passive role. There is a need to note that many learning experiences in schools have now involved pupils in a more active role (e.g., experiential learning, relatively independent individual and group tasks, use of micro-computers), some of which see process itself (rather than content) as the educational experience. The effectiveness of such learning experiences can only be understood by an analysis of the learning tasks and activities as experienced by pupils and by attempts to assess more sophisticated outcomes (e.g., pupil autonomy).

REFERENCES

ANDERSON, L M , EVERTSON, C M., and EMMER, E T (1980) Dimensions in classroom management derived from recent research *J. curric Stud.*, 343-356.

ANDERSON, L. W. (Ed.) (1984). *Time and School Learning*. London: Croom Helm.

BENNETT, S N (1976) *Teaching Styles and Pupil Progress* London Open Books.

BENNETT, S N. (1978). Recent research on teaching· a dream, a belief, and a model. *Br. J educ. Psychol.*, 48, 127-147

BENNETT, S. N., and DESFORGES, C (1983). Recent trends in research on teaching-learning processes. *Education Section Rev* , 7, 4-9.

BENNETT, S. N , DESFORGES, C , COCKBURN, A , and WILKINSON, B. (1984) *The Quality of Pupil Learning Experiences*. London Lawrence Erlbaum.

CALDERHEAD, J. (1984). *Teachers' Classroom Decision-making*. London: Holt, Rinehart and Winston

CENTRA, J. A , and POTTER, D. A (1980) School and teacher effects: an interrelational model. *Rev educ Res* , 50, 273-291

CORNO, L. (1979) Classroom instruction and the matter of time In DUKE, D. L (Ed), *Classroom Management* (78th yearbook of the National Society for the Study of Education) Chicago. NSSE

DOYLE, W. (1983) Academic work. *Rev. educ. Res.*, 53, 159-199

DOYLE, W. (1984). How order is achieved in classrooms: an interim report *J. curric. Stud.*, 16, 259-277

DOYLE, W., and RUTHERFORD, B (1984). Classroom research on matching learning and teaching styles. *Theory into Practice*, 23, 20-25.

EGGLESTON, J F., GALTON, M J , and JONES, M E. (1976) *Processes and Products of Science Teaching*. London: Macmillan.

FITZ-GIBBON, C T , and CLARK, K S (1982). Time variables in classroom research: a study of eight urban secondary school mathematics classes *Br. J. educ. Psychol.*, 52, 301-316.

FREIBERO, H. J (1983). Consistency· the key to classroom management. *J. Educ. for Teaching*, 9, 1-15.

GAMMAGE, P (1982). *Children and Schooling* London· Unwin.

GOOD, T L , and BROPHY, J. E (1980) *Educational Psychology. A Realistic Approach* (2nd edition) New York Holt, Rinehart and Winston

HAERTEL, G D, WALBERG, H J., and WEINSTEIN, T. (1983). Psychological models of educational performance· a theoretical synthesis of constructs. *Rev. educ. Res.*, 53, 75-91

HOWE, M J. A. (1984) *A Teacher's Guide to the Psychology of Learning.* Oxford Blackwell

KOUNIN, J S. (1970). *Discipline and Group Management in Classrooms.* New York Holt, Rinehart and Winston.

KYRIACOU, C (1983). Research on teacher effectiveness in British secondary schools *Br. educ res. J*, 9, 71-80

KYRIACOU, C, and MCKELVEY, J (1984) Individual differences in 'effective' teaching behaviour among secondary school teachers an ESRC Research Grant Report. (Unpublished report, grant ref HR 8791.)

KYRIACOU, C., and NEWSON, G. (1982). Teacher effectiveness a consideration of research problems. *Educ Rev.*, 34, 3-12.

PETERSON, P L., and WALBERG, H. J. (Eds.) (1979). *Research on Teaching.* Berkeley McCutchan.

ROBERTSON, J (1981). *Effective Classroom Control.* London: Hodder and Stoughton

ROSENSHINE, B. (1971) *Teaching Behaviours and Student Achievement.* Slough National Foundation for Educational Research.

RUTTER, M., MAUGHAN, B., MORTIMORE, P., and OUSTON, J (1979) *Fifteen Thousand Hours Secondary Schools and their Effects on Children.* London· Open Books.

RYAN, D. W, and BUTLER, L F. (1982) Effectiveness of teaching the Canadian experience *Stud. in educ. Eval.*, 7, 247-262

TOMLINSON, P (1981). *Understanding Teaching Interactive Educational Psychology* London McGraw-Hill.

WATERHOUSE, P. (1983). *Managing the Learning Process* London· McGraw-Hill.

WINNE, P H., and MARX, R. W (1977) Reconceptualizing research on teaching *J educ. Psychol.*, 69, 668-678.

WOODS, P. (Ed.) (1980a) *Pupil Strategies.* London: Croom Helm

WOODS, P. (Ed.) (1980b). *Teacher Strategies.* London: Croom Helm

WRAGG, E. C. (Ed.) (1984). *Classroom Teaching Skills.* London. Croom Helm.

(*Manuscript received* 4*th October*, 1984)

Br. J. educ. Psychol., 55, 156-163, 1985

THE RELATIONSHIPS BETWEEN TEMPERAMENT CHARACTERISTICS, INTELLIGENCE, TASK-ENGAGEMENT AND MATHEMATICS ACHIEVEMENT

By ZEMIRA R. MEVARECH
(*Bar-Ilan University, Israel*)

SUMMARY This study investigated the relations between pupils' temperament, intellectual ability, Time On Task (TOT) and mathematics achievement. 87 second grade and 104 fourth grade children were observed for three weeks during arithmetic classes. The most important finding to emerge was the strong and consistent relationship between the pupils' temperament characteristics and academic achievement. The final multiple R between temperament, intelligence, Time On Task and mathematics achievement was 0 68 for second graders and 0·56 for fourth graders. Of particular interest was the finding that for the second grade temperament characteristics contributed to the prediction of mathematics achievement more than intellectual skills The addition of behavioural style information to cognitive and task engagement information yielded a more accurate prediction of achievement than that obtained from either alone.

INTRODUCTION

CONSIDERABLE attention has recently focused on children's temperament (for comprehensive reviews see Keogh and Pullis, 1980; Hubert *et al.*, 1982). Research in this field has primarily formulated temperament in terms of stylistic aspects of behaviour with style referring to the "how" of behaviour as opposed to its content (the "what") or motivation (the "why") (Buss and Plomin, 1975; Thomas and Chess, 1977). Studies have shown that *how* children approach, attend, avoid, and respond to school tasks affects adjustment to preschool demands (Carey *et al.*, 1977), social interactions (Billman and McDevitt, 1980) and some behavioural disorders (Keogh and Pullis, 1980). It seems that individual differences in behavioural styles which are evident in infants during the first weeks and months of life are associated with children's behaviour in school settings (Thomas and Chess, 1977). But there is a related question. To what extent does temperament relate to academic achievement?

According to Thomas and Chess (1977), temperament and intelligence are two different attributes that independently affect classroom behaviour. Thomas and Chess assumed that "the meaning of temperament lies not in the child's possession of particular attributes *per se*, but rather in the extent to which the attributes coincide with the contextual demands" (Lerner *et al.*, 1982, p. 150). In school, the desirable repertoire involves high persistence, positive approach, low distractibility and ability to adapt to new situations Thus, even when intelligence is controlled, children possessing such repertoires are expected to be highly engaged on tasks resulting in high achievement gains. In other words, task engagement is hypothesised to mediate pupil characteristics (temperament and intelligence) and academic achievement. Since school demands may show age as well as setting variation, it is quite possible that the association between temperament and achievement will vary across grade levels or across school contexts.

While previous research on school performance has confirmed the linkage between intellectual ability, Time On Task (TOT) and achievement (e.g., McKinney *et al.*, 1975; Anderson, 1976; Bloom, 1976), there is only limited empirical evidence that supports the hypothesised association between behavioural style and achievement (Thomas and Chess, 1977). These results, however, focused only on the differential temperamental repertoires without controlling intelligence or time-on-task. In fact, no previous research has examined the four-variable model linking temperament and intellectual characteristics to on-task behaviour and achievement, although Pullis and Cadwell (1982) indicated that temperament is a crucial variable in school setting.

The purpose of the present study was to investigate this model. The following research questions were addressed: (*a*) What is the role of temperament characteristics in predicting mathematics achievement at different grade levels and to what extent is it influenced by intelligence? (*b*) How much of the variability in achievement scores across age level can be attributed to individual differences in classroom behaviour and how much is accounted for by behavioural styles and intelligence? In order to control for variation in contextual demands all observations and assessments were employed in mathematics classes wherein the instructional method was the lecture-discussion strategy. This method is most widely used in mathematics classrooms (Thompson, 1980).

METHOD

Subjects

Eight classes (N = 191), 87 second grade children and 104 fourth grade children studying in four Israeli public schools comprised the sample for this study. Approximately one half of the pupils were boys (N = 94). Since no significant differences were found between males and females on measures of temperament, intelligence, task engagement and achievement, sex was not discussed beyond this point. The majority of the children came from low to middle class families as assessed by the Israeli Board of Education Scale.

Observational technique

In order to assess task engagement, pupil behaviour was categorised as "on" or "off" task. On-task engagement includes paying attention during classroom activity, working on academic tasks or assignments, discussing concepts learned in the classroom with peers or with the teacher, and reading mathematics problems. Off-task behaviour refers to task-irrelevant activities, visual wandering, daydreaming, aggressive behaviour, conversation with the teacher or peers that is social in nature, and inattention to classroom activities. Following Anderson's study (1976), Time On Task (TOT) was defined as the ratio of "on" task behaviour to the sum of "on" and "off" task responses.

All children in each class were observed for three weeks by trained observers during arithmetic classroom only. The observational procedure was conducted as follows. One child, randomly selected, was observed for 10 seconds and his or her behaviour was coded as "on" or "off" task. Then another pupil was observed in the same way. After the behaviour of all the children was coded, another turn of observation began with the first child and continued through the class period. The TOT measure was based on the same total observation time for each subject (180 observations).

Achievement and ability measures

The academic progress of each subject was measured by the Israeli Board of Education Arithmetic Achievement Test (AAT) which includes mathematical concepts, computations and problems. The alpha reliability scores computed by Kuder-Richardson formula number 21 were 0·93 and 0·81 for second and fourth grades respectively. In addition, the teachers rated the child's level of arithmetic achievement (TR) on a 5-point scale ranging from very poor to excellent. The Advanced Progressive Matrices (Raven, 1960) was used to measure intelligence.

Temperament ratings

Temperament was assessed by the Shortened Teacher Temperament Questionnaire (STTQ) (Thomas and Chess, 1977, p. 22) consisting of nine items each corresponding to one characteristic of temperament as defined by Thomas *et al.* (1968). Since in our study teachers did not have adequate information regarding Rhythmicity, this item was omitted from the questionnaire. Thus, the remaining items assessed the following eight temperament characteristics: Activity Level, Approach, Adaptability, Intensity, Positive Mood, Persistence, Distractibility, and Threshold of Response.

Data reduction procedure

The STTQ was factor-analysed to determine whether the eight dimensions proposed by Thomas and Chess (1977) could be distinguished or whether they were correlated as reported by Pullis and Cadwell (1982). The factor analysis extracted two common factors accounting for 60 per cent of the variance. Table 1 presents the factor pattern after the varimax rotation.

TABLE 1

FACTOR LOADING OF TEMPERAMENT ITEMS (POOLED WITHIN GRADES)

Item No	Dimension	Factor 1 Reactivity	Factor 2 Task Orientation
1	Activity Level	0·731	0·159
2	Approach	0·681	0·295
3	Adaptability	0·423	0·701
4	Threshold	0·433	0·588
5	Intensity	0·807	−0·030
6	Positive Mood	0·578	0·272
7	Distractibility	0·065	−0·510
8	Persistence	0·283	0·525

The first factor, Reactivity (R), contained items from dimensions of Activity Level, Approach, Intensity, and Mood. Pupils scoring high on this subscale tend to react positively to new stimuli, express positive mood, and manifest a high level of activity and energy. The second factor, Task Oriented Behavioural Style (TOBS), was loaded with items from the dimensions of Adaptability, Threshold, Persistence, and Distractibility. This subscale reflects the pupils' ability to persist on tasks until completion, to adapt easily to new situations, to be sensitive to changes and not to be distracted. Averaging the ratings on items with the highest loadings (greater than 0·50) on each factor formed the subscale scores.

RESULTS

The basic statistical method of this study is multiple regression analyses. Two structured regression equations were calculated for each grade level. In the first, pupil scores on an arithmetic achievement test (AAT) served as the criterion, and in the second, teachers' ratings of arithmetic achievement (TR) were used as the dependent variable. In both equations, the predictor variables were the two subscales of temperament, the measure of intelligence, and the amount of time a pupil was engaged on-task divided by the total allocated academic time (TOT). Since the theoretical conceptualisation of this study conceives classroom behaviour as mediating the effects of pupils' characteristics on achievement, our approach was to account for as much variance in mathematics scores as possible by entering TOT first and then finding the amount of additional variation explained by pupil characteristics. In tracing backward from the factor closest to the dependent variable (TOT) to the more remote ones (temperament and intelligence), the model attempts to determine whether resorting to pupils' entry characteristics will result in a meaningful increment in the proportion of variance explained by TOT.

Before examining the above model, the correlations between temperament characteristics, intelligence, TOT and achievement were examined. Table 2 provides the zero order correlations. For ease of comparison between the correlation coefficients of second and fourth grades, the former are reported in the upper half of the matrix and the latter are reported in the lower half of the matrix.

TABLE 2

CORRELATIONS BETWEEN DEPENDENT AND INDEPENDENT VARIABLES

	AAT	TR	TOBS	REACT.	IQ	TOT
AAT	—	62	61	25	38	47
TR	57	—	74	38	49	32
TOBS	33	54	—	55	36	48
REACT	31	35	56	—	31	32
IQ	45	51	28	32	—	11
TOT	40	45	52	28	21	—

Correlations above the diagonal are for second graders and those below the diagonal are for fourth graders Decimal points omitted

Task Oriented Behavioural Style. Table 2 indicates that TOBS was highly related to both measures of achievement. These relationships were slightly influenced by intelligence. Controlling for ability reduced the zero order correlations from 0·55 to 0 48 at second grade, and from 0·32 to 0 25 at fourth grade. Similar results were obtained for the relationships between TOBS and TR. Also here, the zero order correlations and the first order partial correlation (controlling for ability) were very similar (partial $r = 0{\cdot}68$ and 0·48 for second and fourth grades respectively).

Although TOBS was significantly correlated with achievement, the correlation coefficients were not equal at second and fourth grades While TOBS explained 30 per cent of the variance in AAT scores at second grade, it explained approximately 10 per cent of the achievement variance at fourth grade. Using Fisher's procedure for correlation comparisons of independent samples (Glass and Stanley, 1971) indicated that this difference was statistically significant (Fisher's $z = 2{\cdot}03$, $P < 0{\cdot}05$). Attempts to predict TR yielded a similar pattern of relationships as that obtained for AAT scores. The correlations between TOBS and TR were significantly higher at second grade than at fourth grade with and without controlling for ability (Fisher's $z = 2{\cdot}53$, $P < 0\ 05$ for the difference between the zero order correlations and Fisher's $z = 2{\cdot}07$, $P < 0{\cdot}05$ for the difference between the first order partial correlations)

Reactivity. In general, Reactivity was only moderately correlated with AAT scores and TR. It explained, at the most, 15 per cent of the variance in achievement. Partialling out ability reduced the correlations with AAT (from 0·25 to 0 15 at second grade and from 0·31 to 0·20 at fourth grade) as well as with TR (from 0·38 to 0·27 at second grade and from 0·35 to 0 23 at fourth grade). Using Fisher's procedure as before indicated that the relationships between Reactivity and achievement was not significantly different across grade levels.

Intelligence. As was found in previous research, intelligence was highly correlated with achievement (AAT and TR). However, a comparison between TOBS and intelligence showed that at second grade TOBS explained approximately twice the variance in academic achievement than did ability, whereas at fourth grade intelligence explained approximately twice the variance in AAT scores than TOBS. The difference between these correlations (TOBS and achievement *vs.* IQ and achievement) was tested using Hotelling's test for the difference between correlations computed for the same sample (for a description of the procedure, see Guilford, 1956). The difference was marginally significant at second grade ($t(84)=1{\cdot}74$, $P<0{\cdot}07$) and insignificant at fourth grade ($t(101)=1{\cdot}41$, $P>0{\cdot}05$). It is interesting to note that at fourth grade teacher ratings of mathematics achievement were equally related to pupils' intelligence and temperaments whereas at second grade the relationship between intelligence and TR ($r=0{\cdot}49$) was significantly weaker than that between TOBS and TR ($r=0{\cdot}74$) (Fisher's $z=3{\cdot}2$, $P<0{\cdot}05$)

Time On Task. As was expected, children possessing positive TOBS tended to be engaged on tasks even when ability was partialled out (partial $r=0{\cdot}24$ and $0{\cdot}51$ for second and fourth grades respectively). The correlation was significantly stronger for fourth graders than for second graders (Fisher's $z=2{\cdot}12$, $P<0{\cdot}05$). This latter finding suggested that classroom behaviour of older children was more affected by TOBS than that of younger children. In contrast, the correlation between Reactivity and TOT was much the same across grade levels ($r=0{\cdot}30$).

Table 2 further indicates that pupils who were more engaged on tasks scored higher on the achievement test and on TR. Nearly 20 per cent ($r=0{\cdot}47$ and $0{\cdot}40$ at second and fourth grades respectively) of the variability in AAT scores can be explained by TOT. These effects were sustained even when ability was controlled (partial $r=0{\cdot}46$ and $0{\cdot}34$ at second and fourth grades respectively).

Predicting achievement

Given the importance of temperament, intelligence and TOT, how much of the variability in achievement across grade level was explained by all these variables? Table 3 summarises the findings.

The multiple R was equal to 0·68 at second grade and 0·56 at fourth grade. Since the model usually does not fit the population as well as it fits the sample from which it is derived, multiple R was adjusted to fit the model to the population (Norusis, 1983). The adjusted multiple Rs were equal to 0·66 and 0·56 at second and fourth grades respectively, suggesting a good fit of the data to the model.

Further inspection of Table 3 reveals that after TOT was entered to the regression equation, the variation in the measures of pupils' characteristics accounted for an additional 24 per cent of the total variance in second grade and 15 per cent of the variance at fourth grade. At second grade, most of the additional accounted variation is due to temperamental characteristics (20 per cent), whereas in fourth grade the major contributor was the IQ measure, accounting for 14 per cent of the additional variance.

The relative impact of the independent variables on achievement is described by the standardised partial regression coefficients shown in Table 3. At second grade a difference of one standard deviation in TOT or in intelligence was associated with an increase of nearly one-third standard deviation in achievement. A similar pattern of relationship was found in fourth grade. In other words, in order to obtain an improvement of one-third SD in AAT scores (approximately 7 points), an increase of one SD in TOT (approximately 15 per cent) or intelligence is needed. The beta weights also indicate that task orientation had fairly sizable effects at second grade but very small effects at fourth grade

TABLE 3

MULTIPLE REGRESSION ANALYSIS SUMMARY

Predicting Arithmetic Achievement Scores

	Variables Entered	Multiple R	F change	beta
Grade 2	TOT	0 47	23 13*	0 37
	TOBS	0 65	14 14*	0 43
	IQ	0 67	1 49	0 38
	Reactivity	0 68	0 38	−0 13
Grade 4	TOT	0 40	18 86*	0 27
	IQ	0 55	9 90*	0 36
	Reactivity	0 56	0 48	0·10
	TOBS	0 56	0 0	0 05

Teacher Ratings of Arithmetic Achievement

	Variables Entered	Multiple R	F change	beta
Grade 2	TOT	0·32	9 10*	0 13
	TOBS	0 74	42 86*	0 62
	IQ	0 78	3 50*	0·24
	Reactivity	0 78	0 0	0·01
Grade 4	TOT	0 45	24 75*	0·18
	IQ	0 61	14 37*	0 40
	TOBS	0 68	4 89*	0 33
	Reactivity	0 68	0·0	0 0

* P<0 5

DISCUSSION

The findings of this study lend support to the assumption that temperament is an important determinant of young children's academic achievement. Results showed that at second grade TOBS explained more than twice the variance in mathematics achievement than did a cognitive test, and that the variance that was contributed by these two measures was largely independent of that provided by TOT measures. Thus, the addition of temperament information to cognitive and task-engagement information yielded a more accurate prediction of achievement than could be obtained from either alone.

The data further show two grade-related trends First, the relationship between TOBS and TOT was weaker for younger children than for older ones While seemingly surprising, other researchers have reported similar results. Thomas *et al.* (1968, 1977) and Pullis and Cadwell (1982) found that extremely adaptable children exhibited more often adjustment difficulties during the first years of school which gradually decreased as a function of grade level. Hallahan and Reeve (1980) indicated that younger children had more difficulties in attending to relevant stimuli in the presence of proximal irrelevant stimuli than older children. Since the present study used a composite score of behavioural styles it is impossible to know which specific components of temperament were critical to task engagement and achievement. Future research may address this issue.

The other grade related trend concerned the relationship between TOBS and achievement at second grade It is possible that younger children were more affected by their temperaments because they had not yet developed attentional skills, while older children had adapted more to school demands (Buss and Plomin, 1975; Hallahan and Reeve, 1980). If this is indeed the case, an explanation may be suggested for the small impact of intelligence on achievement when TOT and TOBS was controlled. It is widely assumed that various intelligence tests require attentional skills including persistence, low distractibility, adaptability and threshold. Since there is a clear developmental trend in children's ability to sustain attention (Hallahan and Reeve, 1980), it is likely that the effects of IQ on achievement at second grade is mainly due to selective attention and low distractibility. As children get older, they are less influenced by distractions and thus they are more affected by other dimensions of intelligence. Further research might reveal the pattern of change in the relationships between temperament and achievement as a function of the child's age.

While the present study provided supporting evidence for the association between temperament, intelligence, TOT and achievement, its generalisation is affected by several conceptual and methodological limitations. On a conceptual level it is necessary to investigate further the number and independence of the temperament dimensions. This issue is still an open question in the psychological literature in which the proposed number of dimensions are three (Pullis and Cadwell, 1982), four (Buss and Plomin, 1975), and nine (Thomas and Chess, 1977). On a methodological level, the study is limited by the use of the STTQ in which each temperament dimension is represented by only one item. Although there is little evidence to demonstrate the validity and reliability of this instrument, the fact that the factors obtained were similar to those reported by Pullis and Cadwell (1982) shows a high construct validity

Keeping in mind the problems of measurement, this study has several implications. First, the study pinpoints the importance of temperament characteristics in school situations. Teachers may be trained to use both cognitive as well as temperament information in planning classroom activities Second, the study suggests that altering the pupils' classroom behaviour may have a positive effect on academic achievement Since intelligence and temperamental characteristics are relatively stable variables (Buss and Plomin, 1975; Bloom, 1981), modifying pupils' TOT might be easier to implement and beneficial for increasing school outcomes Finally, the present study used regression methods and thus it was impossible to infer causality. It would be useful to examine such trends by conducting experimental studies. The challenge for future research on temperament characteristics and pupil achievement will be to understand the mechanism through which variations in behavioural styles affect school performance.

REFERENCES

ANDERSON, L W (1976) An empirical investigation of individual differences in Time On Task. *J educ Psychol*, 68, 226-233

BILLMAN, J , and MCDEVITT, S C (1980) Convergence of parental and observer ratings of temperament with observations of peer interaction in nursery school *Child Dev* , 51, 395-400

BLOOM, B S (1976) *Human Characteristics and School Learning* New York McGraw-Hill

BLOOM, B S. (1981) *All Our Children Learning* New York McGraw-Hill.

BUSS, A. H , and PLOMIN, R (1975) *A Temperament Theory of Personality Development* New York: Wiley.

CAREY, W , FOX, M , and MCDEVITT, S (1977) Temperament as a factor in early school adjustment *Pediatrics*, 60, 621-624,

GLASS, G V., and STANLEY, J C (1971) *Statistical Methods in Education and Psychology* Englewood Cliffs, N.J.: Prentice Hall

GUILFORD, J P (1956) *Fundamental Statistics in Psychology and Education* New York McGraw-Hill

HALLAHAN, D P., and REEVE, R E (1980) Selective attention and distractibility In KEOGH, B K (Ed.), *Advances in Special Education* Greenwich, Conn. JAI Press

HUBERT, N C., WACHS, T D., PETERS-MARTIN, P , and GANDOUR, M. J (1982) The study of early temperament: measurement and conceptual issues *Child Dev* , 53, 571-601

KEOGH, B., and PULLIS, M. (1980). Temperament influences on the development of exceptional children In KEOGH, B K. (Ed), *Advances in Special Education*. Greenwich, Conn : JAI Press

LERNER, R. M., PALERMO, M., SPIRO III, A , and NESSELROADE, J R (1982) Assessing the dimensions of temperamental individuality across the life span the dimensions of temperament survey (DOTS). *Child Dev* , 53, 149-159

McKINNEY, J., MASON, J , PERKERSON, K , and CLIFFORD, M (1975). Relationship between classroom behavior and academic achievement. *J. educ Psychol.*, 67, 198-203

NORUSIS, M. J (1983). *SPSS-x Introductory Statistics Guide*. New York McGraw-Hill.

PULLIS, M , and CADWELL, J. (1982). The influence of children's temperament characteristics on teachers' decision strategies. *Am educ Res J* , 19, 165-182.

RAVEN, J S (1960) *Progressive Matrices* London· H K Lewis.

THOMAS, A , CHESS, S , and BIRCH, H G (1968) *Temperament and Behavioral Disorders in Children* New York New York University Press

THOMAS, A , and CHESS, S. (1977) *Temperament and Development* New York Bruner and Mazel

THOMPSON, S B (1980) Do individualized mastery and traditional instructional systems yield different course effects in college calculus? *Am educ Res J* , 17, 361-375

(*Manuscript received 12th September*, 1984)

RESEARCH NOTES

Br. J. educ. Psychol., 55, 164-166, 1985

THE TRANSCENDENTAL MEDITATION TECHNIQUE AND POSTGRADUATE ACADEMIC PERFORMANCE

BY PAUL KEMBER

(*Department of Ergonomics, Cranfield Institute of Technology, Bedford*)

SUMMARY. Twenty students taking the first year of an MSc course at Cranfield Institute of Technology were randomly allocated to a Transcendental Meditation or control group. There was some evidence that six months' practice of Transcendental Meditation produced an improvement in postgraduate academic performance

INTRODUCTION

A number of research studies indicate that the Transcendental Meditation (TM) technique may produce a reduction in stress and an improvement in the functional organisation of the nervous system (e.g., Dillbeck and Bronson, 1981; Jevning *et al.*, 1983).

Dillbeck *et al.* (1979) suggest that these changes may have important applications in education, and summarise several studies indicating positive cognitive and affective effects of Transcendental Meditation in secondary and undergraduate students.

The present pilot study is the first experiment to examine the effects of Transcendental Meditation on the academic performance of postgraduate students.

METHOD

Subjects

The subjects were 20 students taking the first year of the MSc course in Industrial Engineering and Production Management at the Cranfield Institute of Technology. All 10 students in the experimental group completed the course, but only six of the control group sat all the examinations (three transferred to another course and a fourth was convalescing from an accident).

All the students included in the analysis were males. The mean age was 26·0 years for the experimental group and 25 0 years for the control group.

Design and measurements

The subjects were randomly allocated to two groups. Group I received instruction in Transcendental Meditation, as taught by Maharishi Mahesh Yogi, and Group II acted as a control group. The instruction took place in January, 1982. Neither group was aware that the purpose of the experiment was to assess the effects of Transcendental Meditation on academic performance Together these conditions therefore controlled for artefactual effects of expectation, motivation, and self-selection.

All the examination results standard to the MSc course were assessed in December, 1981 (pretest) and June 1982 (post-test). Nine examinations were taken at pretest and 12 at post-test.

None of the staff involved in examination assessment was aware of group membership

ANALYSIS AND RESULTS

The overall mean for all the subjects' performance on all exams was 63·4 at pretest (SD 28·7) for the experimental group and 63 6 for the control group (SD 29·5); t = 0·03, df = 141, P > 0·1. At post-test the experimental group had a mean score of 61·3 (SD 20·5), compared with 59·0 (SD 21·4) for the control group; t = 0·74, df = 190, P > 0·1.

TABLE 1

MEAN SCORES FOR EACH GROUP ON ALL EXAMINATIONS

Pretest

	Ergonomics	Prod Control	Fortran	Maths	Network Analysis	Computing	Group Tech	Eng Drawing	Machining Ops
Experimental	50·3	59·0	59·3	52·1	61·2	131·8	60·1	42·5	45·0
Control	49·4	59·7	56·5	57·7	60·8	130·7	62·3	45·0	47·7

Post-test

	Ergonomics	Prod Control	Fortran	Maths	Network Analysis	Computing	Group Tech	Statis-tics	Manu Systems	Work Org	General Manu	Communi-cations
Experimental	49·4	48·3	60·0	112·6	58·9	54·5	72·8	56·3	60·0	56·4	45·7	60·1
Control	48·6	42·1	56·5	117·0	53·3	53·2	71·7	63·0	51·4	56·0	43·3	57·8

The mean score of all the students for each examination was then calculated for the two groups (Table 1). At pretest the mean score of the experimental group was greater than that of the control group in 4 of the 9 subjects (44 4 per cent), while at post-test the experimental group obtained a higher average in 10 of the 12 subjects (83·3 per cent). The post-test difference was statistically significant ($P = 0{\cdot}019$, binomial test).

When standard scores were considered, there were no significant differences between the groups at pretest, but at post-test significantly more standard scores for the experimental group were greater than two standard deviations above the mean. (Five scores for the experimental group, compared with none for the control group; $P = 0{\cdot}03$, Fisher Exact)

DISCUSSION

The results of this pilot study are consistent with previous work indicating that improvements in academic performance can result from the practice of Transcendental Meditation, and support the suggestion that the applications of the Transcendental Meditation technique to education should be studied further.

Acknowledgments —The author gratefully acknowledges the support given for this pilot study by Professor Brian Hundy, Mr. Paul Galgut and The School of Production Studies.

Requests for reprints should be addressed to· Paul Kember, Department of Ergonomics, School of Production Studies, Cranfield Institute of Technology, Cranfield, Bedford, MK43 0AL

REFERENCES

Dillbeck, M. C., and Bronson, E C. (1981). Short-term longitudinal effects of the Transcendental Meditation technique on EEG power and coherence. *Int J Neuroscience*, 14, 147-151.

Dillbeck, M. C , Aron, A , and Dillbeck, S. (1979) The Transcendental Meditation program as an educational technology. *Educ Technol.*, 19, 7-13

Jevning, R., Wilson, A F., O'Halloran, J P , and Walsh, R. N (1983) Forearm blood flow and metabolism during stylized and unstylized states of decreased activation. *Am J Physiology*, 245, R110-R116

(*Manuscript received 7th October*, 1984)

Br. J. educ. Psychol., 55, 167-169, 1985

LAWSEQ: ITS RELATION TO OTHER MEASURES OF SELF-ESTEEM AND ACADEMIC ABILITY

BY J. G. HART

(*College of St Paul and St. Mary, Cheltenham*)

SUMMARY. This research was concerned with two issues an examination of the validity of the questionnaire LAWSEQ (Lawrence, 1981), and empirical verification for the generally held assumption that there is a relationship between self-esteem and levels of anxiety, academic self-image, academic achievement and ability. The tests were administered to 128 fourth-year junior school children in the term prior to their transfer to secondary school. LAWSEQ was found to be reasonably stable over a period of four months ($r_{tt} = 0·64$) and had a fairly high correlation ($r = 0·73$) with the Coopersmith Self-Esteem Inventory (SEI) (Coopersmith, 1967) Significant correlations were detected between self-esteem, levels of anxiety and academic self-image No significant relationship was found between self-esteem and academic achievement or verbal reasoning

INTRODUCTION

Lawrence (1981) defined self-esteem as the "child's effective evaluation of the sum of his or her characteristics both mental and physical", and devised a questionnaire (LAWSEQ) that purports to provide a measure of self-esteem. Lawrence provided some information on the reliability and validity of this scale but he recognised the need for further evaluative work to be undertaken

Self-esteem has been related to a wide variety of psychometric measures at one time or another, for example, educational achievement, anxiety, responses to threat or stress, teachers' attitudes and maladjustment and delinquency (Lawrence, 1973; McKinney *et al.*, 1978). There is quite a body of research reporting positive correlations between self-esteem and school achievement (e.g., Piers and Harris, 1964, Purkey, 1970), although later studies have not produced consistent results (e.g., Thomas, 1973)

It is well recognised that the self-esteem of the child in school is a crucial factor in his/her learning at all levels of success (Shavelson and Bolus, 1982; Marsh *et al.*, 1983). Some believe that it has a predictive value for academic achievement equal to, or better than, measures of intelligence (Gammage, 1982). All the evidence points to well-adjusted children evaluating themselves positively (Thompson, 1974) and therefore many think that one of the main goals of schooling should be to assist pupils in this direction.

It would appear useful, therefore, to gain some further empirical verification for these beliefs, and to obtain some validity for a recent measure of self-esteem, such as LAWSEQ.

METHOD

The sample

A total of 65 girls and 63 boys, aged between 10 years 6 months and 11 years 5 months (in their final year at junior school) comprised the sample. They attended two schools in a semi-rural Midlands area, both of which fed two co-educational comprehensive schools Children identified as having special educational needs were excluded from the sample.

Variables

The instruments chosen for self-esteem measurement were LAWSEQ (Lawrence, 1981) and the Coopersmith Self Esteem Inventory (Coopersmith, 1967) The full SEI has test/retest reliabilities ranging from 0·88 over a 5-week period to 0·70 over a 3-year period (Coopersmith, 1967). In this study a shortened version of 25 questions was used which is reported to correlate at 0·86 with the 58-item version (Argyle and Lee, 1972).

The other variables selected as tests which are likely to be used in conjunction with LAWSEQ were:

(1) National Foundation for Educational Research Test D3 English (NFER, 1976)
(2) Test DE1 Mathematics (NFER, 1970)
(3) Moray House Test 95 Verbal Reasoning (University of Edinburgh, 1977)
(4) Academic Self-Image Scale (Barker Lunn, 1970)
(5) The Boxall Test of School Anxiety (Boxall, 1961) modified by Dale (1969)

Procedure

The Moray House VRQ is part of the Education Authority's assessment for fourth-year junior pupils, and was administered in the first term of the year.

The NFER tests of Maths and English were given by the schools as part of their annual assessment procedure during the summer term

The first LAWSEQ was administered in February, and this was repeated four months later, together with the Coopersmith SEI, Academic Self-Image and the Boxall Test of School Anxiety.

RESULTS

Table 1 gives a Pearson correlation matrix for all variables.

LAWSEQ proved to be reasonably stable over time ($r = 0{\cdot}64$) and correlated highly with the Coopersmith SEI ($r = 0{\cdot}72$, $N = 122$, $P < 0{\cdot}001$). A high positive correlation was also found between self-esteem and academic self-image ($r = 0{\cdot}39$, $N = 122$, $P < 0{\cdot}001$). A high negative correlation was found between self-esteem and anxiety ($r = -0{\cdot}55$, $N = 122$, $P < 0{\cdot}001$). Correlations between self-esteem, academic achievement and ability in terms of Mathematics, English and Verbal Reasoning were not found to be significant, although academic self-image and anxiety both correlated significantly at the $P < 0{\cdot}05$ level with all three tests of academic ability.

Internal reliability coefficients

The internal reliability was also obtained for the first testing of LAWSEQ. The reliability coefficients for the 16 items were: Alpha = 0·616, Standardised item Alpha = 0·597

TABLE 1

PEARSON INTER-CORRELATION COEFFICIENTS

	LAWSEQ 1st Testing	LAWSEQ 2nd Testing	Coopersmith SEI	Academic Self-Image	Boxall Anxiety	VRQ	NFER Maths
LAWSEQ 2nd Testing	0·64 ***						
Coopersmith SEI	0·49 ***	0·73 ***					
Academic Self-Image	0·30 **	0·39 ***	0·42 ***				
Boxall Anxiety	−0·49 ***	−0·55 ***	−0·52 ***	−0·43 ***			
VRQ	0·12	0·06	0·10	0·58 ***	−0·27		
NFER Maths	0·18 *	0·10	0·17	0·51 ***	−0·29 *	0·80 ***	
NFER English	0·06	0·10	0·20	0·56 ***	−0·15	0·80 ***	0·75 ***

* $P<0{\cdot}05$ ** $P<0{\cdot}01$ *** $P<0{\cdot}001$

DISCUSSION AND CONCLUSIONS

LAWSEQ proved to be relatively stable and consistent over a period of four months and correlated highly (r = 0·73) with another well-established measure, the Coopersmith SEI. The high correlation between low self-esteem and high anxiety supports other consistent findings. The relationship between the two is interesting. It may be that high anxiety produces low self-esteem or vice versa, but it is also possible that anxiety is a subset of self-esteem and that the instruments for measuring both ask similar types of question.

Self-esteem was not found to be a predictor of academic achievement whereas the highest correlations were, perhaps as expected, between Verbal Reasoning and Maths and English scores. Academic self-image was found to have a higher relationship with both scores of ability and academic achievement than self-esteem.

LAWSEQ was devised to assist in the identification of children who may suffer from poor self-esteem, and as such would appear to achieve its aims, but its predictive value for academic performance would seem to be questionable.

ACKNOWLEDGMENT. — I should like to thank Dr. Colin Terrell, College of St. Paul and St. Mary, Cheltenham, and Professor Philip Gammage, School of Education, University of Nottingham, for their supervisory advice. I should also like to thank the children and staff of the schools in the sample, who made the research possible.

REFERENCES

ARGYLE, M , and LEE, V. (1972). *Social Relationships.* Milton Keynes: Open University Press.
BARKER LUNN, J. C. (1970). *Streaming in the Primary School.* Slough: NFER.
BOXALL, J. (1961). A study of some relationships between anxiety and failure in learning to read. Dissertation for the Diploma in Child Psychology, University of Birmingham
COOPERSMITH, S. (1967). *The Antecedents of Self-esteem.* San Francisco: W. H. Freeman.
DALE, R. R. (1969). Anxiety about school among first year grammar school pupils and its relation to occupational class and co-education. *Br. J. educ. Psychol.*, 39, 18-26.
GAMMAGE, P. (1982). *Children and Schooling: Issues in Childhood Socialisation.* Allen and Unwin.
LAWRENCE, D. (1973). *Improved Reading Through Counselling.* London: Ward Lock.
LAWRENCE, D. (1981). Development of a self-esteem questionnaire. *Br J. educ. Psychol*, 51, 245-251.
MARSH, W. H.. SMITH, I. D., BARNES, J., and BUTLER, S. (1983) Self-concept: reliability, stability, dimensionality, validity and the measurement of change. *J. educ. Psychol.*, 75, 772, 790.
MCKINNEY, F., MILLER, D. J., BEIR, I , and BOHANNA, S. R. (1978). Self-concept, delinquency and positive peer culture. *Criminology*, 15, 529-538
PIERS, E., and HARRIS, P. (1964). Age and other correlates of self-concept in children. *J. educ. Psychol.*, 55, 91-95.
PURKEY, W W., (1970). *Self-Concept and School Achievement.* New York: Prentice Hall.
SHAVELSON, R. J., and BOLUS, R. (1982). Self-concept· the interplay of theory and methods. *J. educ Psychol*, 74, 3-17.
THOMAS, J. B. (1973). *Self-concept in Psychology and Education· a Review of the Research.* Slough: NFER.
THOMPSON, B. (1974). Self-concepts among secondary school pupils. *Educ. Res.*, 17, 41-47.

(*Manuscript received* 26*th January*, 1985)

BOOK REVIEWS

BENNETT, N., DESFORGES, C., COCKBURN, A., and WILKINSON, B. (1984). *The Quality of Pupil Learning Experiences.* London: Lawrence Erlbaum, N.J., pp. ix+262, £4·95 pb.

STONES, E. (1984). *Supervision in Teacher Education.* London: Methuen, pp. ix+163, £4·50.

In an ideal classroom all pupils would be engaged in tasks perfectly appropriate to their previous experience, current level of competence and future programme, and teachers would have an impeccable understanding of major factors affecting each pupil's ability to learn. The state of affairs in the real world is, alas, well short of this paragon, and Bennett and his collaborators have conducted a most valuable enquiry into how classroom assignments are conceived and implemented, as well as the extent to which they match the abilities of the children.

The Quality of Pupil Learning Experiences describes a clutch of related studies of tasks set during language and number work in classes containing 6- or 7-year-olds It is far from easy to monitor pupils' responses to their task, as many classroom assignments cannot readily be placed on some linear scale of difficulty and children can be adept at concealing their true ability. Nevertheless the authors have conducted this research with great care and a sensitive recognition of the limitations of what they have found.

The principal matching study involved trained classroom observers keeping records on six target pupils, with equal numbers of boys and girls, above average, average and below average children, in each of sixteen classes. Teachers were interviewed beforehand to elicit the purpose and reasoning behind each task as well as any anticipated problems, and subsequently to ascertain reactions and future plans. In between the interviews, the investigators would observe the pupils' responses and interview them about their understanding of what they were doing.

At the same time they tested each pupil's ability to cope with items which were easier or harder than the task in hand, simple enough, perhaps, when rudimentary operations with one or two digit numbers are involved, but more problematic in certain kinds of language work There were also studies of children after transfer to the junior school, of group work, and an experimental in-service course with experienced teachers.

The researchers found that the four rules of number figured frequently in the tasks set by teachers, and there was extensive practice of routine calculations, as well as neglect of practical relevance, points which have been observed in several HMI reports including the Primary School Survey, giving the lie to those who believe that basic skills in mathematics are being neglected. What is more significant in the matching study, however, is the authors' conclusion that neither bright pupils nor slow learners were adequately provided for.

In language work most tasks involved writing in various contexts and for several purposes, and reading had similar prominence, though only 2 per cent of tasks focused exclusively on reading. Teachers' intentions were stated in terms of procedure rather than cognitive gain, though the authors may, I suspect, have underestimated the implicit nature of the latter. In general the match for both number and language tasks was only estimated at about 40 per cent. A third of tasks were said to be too difficult and a quarter too easy. High attainers were frequently underestimated and low attainers overestimated It was as if teachers were constrained against shifting their expectations too far from the class mean. The junior school study revealed even less of a match, only 30 per cent of number tasks being judged appropriate and high attainers being underestimated in three quarters of all tasks set

Despite reservations one might have about the accuracy of assessing match on a variety of tasks, these findings by Bennett and his colleagues deserve the widest dissemination and scrutiny. The ability to diagnose and then invent lies at the very heart of skilful teaching. This book contains a great deal of case account from the researchers' field notes, which not only amplifies the quantitative data but provides valuable training material for pre- and in-service courses for teachers. Indeed the authors themselves describe, in chapter 9, an in-service course they organised in which teachers learned to analyse descriptions of classroom events as well as to conduct diagnostic interviews with pupils This is the positive way ahead if the team's perceptive enquiry is to be translated into sensitive practice, and ultimately lead to an improvement in teachers' general professional competence.

It is the sharpening of professional skill which is the central theme of Stones' book *Supervision in Teacher Education*, in which he takes forward arguments about the nature of teaching and teacher training which will be familiar to readers of his earlier books on *Teaching Practice* (1972) and *Psychopedagogy* (1979).

The author's principal contention is that, in traditional teacher training courses, supervisors lecture about teaching and then pursue their students around schools during teaching practice uttering platitudes not based on any sound body of psychological theory. In Stones' model of training the supervisor and novice are involved with the pupils themselves in a form of perpetual enquiry. The pupils are seeking the skills they need to solve problems, student teachers are testing their own effectiveness in lubricating this process, and supervisors likewise appraise their own competence at facilitating the whole show.

He writes about teaching problem-solving, which involves preactive, interactive and evaluative phases as the teacher analyses the task, explains the problem, provides feedback and moves on to new problems, indeed all the points brought out in the study by Bennett and his team. He also scrutinises the teaching of concepts, saying that teachers too frequently teach words rather than actual concepts. Supervisors should, he argues, provide beginners with a graded sequence of examples from classroom life, using videotapes and transcripts, which reduce the general chaos and bring clarity to the process of teaching and learning.

The outcome of this exploratory style of supervision is that new teachers are held back from the hurly-burly of classroom life for quite a time. The supervisor spends much more time on in-depth analysis of a videotape of the student's first hesitant steps. One hopes that at the end of it novices are more confident and more perceptive about classroom processes. I have my doubts, and too long a delay can be as sapping of confidence and as generating of anxiety as premature exposure, but Stones' notions offer a well-reasoned challenge to much current practice.

TED WRAGG

BURGESS, R. G., (Ed.) (1984). *The Research Process in Educational Settings: Ten Case Studies*. London: Falmer Press, pp. 282, £14·00 hb, £7·95 pb.

Case studies, reports and papers which emanate from the qualitative field of educational research rarely offer more than a partial reflection of research processes The ten autobiographical accounts which comprise the bulk of material in these pages are concerned with mapping aspects of these processes in predominantly qualitative programmes, projects and individual studies. Arising from recent increased interest and activity in ethnographic-style research, the concern in this book is with the in-field complexities of conducting research and evaluation. The attempt is to communicate that process through first person accounts.

Consequently, what results is a fairly disparate collection of reflective, biographical evidence which promises the reader a share in the kind of insights which are normally reserved for off-the-record discussions between researchers, in the pub. In some cases the promise is undoubtedly fulfilled. In others the looking glass remains unbreached. The two accounts which generate most involvement in the reader are by Martyn Hammersley and David Jenkins. In Hammersley's "The Researcher Exposed: A Natural History" there is a freshness and frankness in the way that he uncovers the uncertainties, happenchances and largely undesigned progress of the researcher in his novitiate Although others deal with later, experienced periods of their research careers, there are only the occasional hints at the dissonance between planning an implementation in naturalistic research and generally they do not have the immediacy of this section The vagaries of research, although echoed in each of the sections, are most upfront in this one: here, and in David Jenkins' account of the politicking in Northern Ireland over the evaluation report he helped write called 'Chocolate Cream Soldiers. Sponsorship, Ethnography and Sectarianism.' This section is the only one which fully recognises the political context of research and evaluation. Who owns the data, at what point and for how long, are central questions for a form of research which relies so heavily upon the vulnerability of participants when they communicate their understandings of the world to the researcher

These contributions were developed from inputs to a workshop in London over two years ago. The research upon which the chapters depend may have taken anything up to ten years prior to that. In a sense this makes the writing more complacent and secure as the writers' present invulnerability and professional standing remain relatively unthreatened. It is one of the endemic problems of the ethnographic form, with its cumbersome mountains of data, its principles and procedures, its time-consuming fieldwork and its progressive search for alternative theoretical understandings, that there develops a growing temporal alienation between the written products of research and the people and events which were its original focus.

However, despite the faint fustiness which clings to the volume and which owes to the passing of time and topics of contention, the collection has much to recommend it. It provides some good sources for those interested in following up the effects of biography upon research, the on-the-ground difficulties of negotiation and the logistics of changing designs. It may be a stimulating reader for those about to embark upon their first ethnographic project and who must realise that failures and reverses are commonplace but not necessarily terminal. It is generally well-orchestrated by Robert Burgess, the editor, whose opening and closing sections begin to set out an important case for the deprivatisation of the research process. His concluding remarks are·

> "Is it a crime if as researchers we do not address the social, ethical and political problems that surround the research process in the study of educational settings?"

If this collection commits a crime then it is the crime of negligence. It says much but it does not reveal enough by half. As Burgess himself points out, there is a great need for the research process to be opened up to the discourse Democracy in the researcher-participant relationship can only result from the researcher sacrificing professional distance for a daily and more vulnerable accountability. Sociologists of education, some of whom are represented in this volume, have a long way to go in this respect. Educational researchers and evaluators from both sides of the Atlantic, who are under-represented in the references in these chapters, have been actively engaged with improving research democracy for the last ten years. The key is in language. Whose language is it? If it is the researcher's then there is no democracy. It would have given this volume a more contentious topicality if these moves to democratise and demystify the research process had been a central concern.

JACK SANGER

CELL, E. (1984). *Learning to Learn from Experience*. Albany, N.Y.: State University Press, pp. xiii+245, n.p.

It is the case with books, as with films, or anything for that matter, that high expectations are the progenitors of disappointment. Such was my feeling about this book, certainly, for I jumped at the title, imagined it to herald just the work I have been looking for, and so set myself up to be let down Having put it aside for a fortnight, I hope I can now be fairer to it

Learning to Learn from Experience presents for us a subject that is long overdue for rediscovery by students of human learning and human education. While our colleagues in animal labs have kept their studies going, the rush of blood to the head caused by the permissive era (and aura) of cognitive psychology has concomitantly drained it from the extremities. We have been preoccupied with the "heady" pursuits of talking and thinking and reasoning and have ignored the mass of experiential learning on which these pinnacles of erudition are perched. Psychologists of education are inclined to make this mistake twice over, being interested in an institution that is itself obsessed with intellectual performance.

But this book promises not just to reattach our sense-ability and response-ability to our think-ability, but to deal with the expansion, the amplification of those abilities, an issue that for me at any rate is at the very heart of what schools should be about And in part the promise is fulfilled. The book does make a significant contribution to the new integrative spirit that is attempting to combine behavioural, cognitive and humanistic ingredients into a coherent and comprehensive account of personal growth. Somewhere between the exaggerated rationalism of "cognition" and the pious world of "Know Thyself" paperbacks is emerging a psychology that has the rigorous concern with process of the former without its limitations, and the importance of the latter without its evangelism

The major problem with the book is that it hasn't got the balance right. Sometimes the language is too heavy, certainly for education students, while at others it is too sickly by half. The second chapter, for example, tries to present some "serious" psychological distinctions between different kinds of learning that are primitive in the extreme, and which (perhaps wisely) are not carried through in the rest of the book. It is as if Cell tried to reinvent the wheel, came out with something with a few corners, and then mislaid it. So we are not offered a great deal of value *qua* psychology. On the other hand there is rather too liberal (in both senses) a sprinkling of rhetoric. Autonomy, enrichment, freedom, creativity, courage, real, authentically human ... Hurray! Defence, invalidate, manipulate, suppression, spiritual bankruptcy... Boo!

Yet although I don't like the layout of the garden, and some of the blooms are too garish, there is a lot of good growth, particularly in the examples and the detail. Some of it is higgledy-piggledy or overcrowded or repetitive, but there is plenty here to interest the reader, and pages in which you can see yourself reflected and clarified. The chapters on "resistance to learning", "changing our emotions", and "learning to know other persons" are particularly valuable. The final section of the book covers "strategies and tools" for experiential learning, some of which is again interesting though annoyingly disorganised, and concludes with a disproportionate discussion of the personal journal or diary as such a tool.

Some Masters' students will enjoy it and benefit from it, and there is certainly something there for the trade, provided one is not too incensed by loose talk. But overall, I suspect, the reaction will be the ambivalent one that Paul Daniels anticipates for his magic: "You'll like it... Not a lot, but you'll like it."

GUY CLAXTON

DAVIS, R. B. (1984). *Learning Mathematics: The Cognitive Science Approach to Mathematics Education.* London: Croom Helm, pp. 392, £22·50.

Cognitive science is currently much in vogue, for reasons which are not hard to find. The boom in information technology has resulted in a frenetic international race (in which the USA and Japan are front-runners, with Western Europe some way behind) to produce and market intelligent artificial systems. Many people — including the Alvey Committee (1982) — believe that the winners of this race will have a position of unprecedented power in the world economy. The design of these artificial systems will increasingly require a theory of intelligence which will analyse functions such as knowing, thinking and understanding in terms that can readily be modelled by a computer. The obvious area in which to look for such a theory is the interface between psychology, computer science, linguistics and epistemology which has become known as cognitive science.

Whether cognitive science can deliver what many fervently believe it can is very much open to debate What is less debatable is that we will be increasingly flooded with cognitive science approaches to a vast range of topics, as more and more researchers — whether from genuine belief or from financial expediency — jump on the bandwagon. Mathematics education is no exception: indeed it has already received more attention than most from cognitive scientists. We should welcome many aspects of this attention, such as the emphasis placed on studying the real behaviour of children attempting mathematics problems, the focus on errors as sources of insight into the learning process, and the development of precise and testable models of mathematical thinking. Brown and Burton's (1978) modelling of children's mathematical errors is a classic example of the potential of this approach. Less welcome is the assumption that all human cognition can be explained purely in information-processing terms, or the belief that the models produced so far are of much practical use in the classroom.

There is a definite need for a book which sets out clearly and dispassionately the strengths and weaknesses of the cognitive science approach to mathematics education. Unfortunately, Davis does not fill the bill. According to its author, the book started as a report on the methods and findings of a seven-year study into mathematics learning, but broadened to take a cognitive science approach to the dichotomy between rote-learning and understanding. At the end of the book, however, I wasn't much wiser about the findings of the long-term study, although I had read some fascinating examples of students' attempts to solve mathematical problems. Nor was I any clearer as to what cognitive science had to say about the rote-learning/understanding split. Indeed, almost all the author's final recommendations to mathematics educators — such as to speed up the pace of the primary curriculum — seemed to owe less to cognitive science than to his own opinions.

CALCUTTA

Learning Mathematics is in fact a rather laboured and *ad hoc* collection of attempts to explain interesting examples of children's mathematical performance in terms of information-processing concepts (procedures, frames, assembly, retrieval and so on). In some cases the explanations are illuminating (as in Friend's (1979) account of addition errors in Nicaraguan children), but in many others they are not. I was, for example, intrigued by a boy (studied by Erlwanger, 1973) who, when asked to write 2/10 and 27/15 as decimals, wrote 1·2 and 4·2 respectively — but I didn't feel that I gained much from being told that in his cognitive apparatus "a super-procedure was selecting the wrong sub-procedure". I was also very disappointed by the way in which the notion of "frame" was invoked to explain a wide variety of phenomena: virtually any error can be explained as the child selecting an inappropriate frame, but that doesn't take you very much further.

Finally, Davis hopes his book "will be of interest both to researchers in the area of cognitive studies, and also to teachers and those who are studying to become teachers". I can't see it helping either group. Researchers will want a much shorter presentation of his actual research findings, while teachers and students will want something clearer, more critical and much cheaper than this.

MARTIN HUGHES

REFERENCES

ALVEY COMMITTEE (1982). *A Programme for Advanced Information Technology.* London· HMSO.

BROWN, J. S., and BURTON, R. R. (1978). Diagnostic models for procedural bugs in basic mathematical skills. *Cognit Sci*, 2 (2), 155-192.

ERLWANGER, S. H. (1973) Benny's conception of rules and answers in IPI mathematics. *J. Children's math. Behav*, 1 (2), 7-26.

FRIEND, J. E (1979) Column addition skills. *J Children's math. Behav*, 2 (2), 29-57.

DUNHAM, J. (1984). *Stress in Teaching.* London: Croom Helm, pp. 176, £12·95.

This book presents a fascinating insight into teacher stress. Jack Dunham has for many years been involved in workshops on teacher stress and has been able to use many of the accounts provided by teachers of their stress experiences to illustrate the main issues warranting attention.

The literature on teacher stress has to some extent neglected a consideration of the pressures concerned with senior management and staff development issues, and this book can be commended for dealing with these areas in a very useful way

I found two issues particularly illuminating. Firstly, the problems for teachers stemming from new demands made upon them. Secondly, the problems experienced by teachers in dealing with colleagues.

In his treatment of coping with stress, I thought the four categories of resources presented (personal, interpersonal, organisational and community) were very helpful. The elaboration of personal and interpersonal resources perhaps erred overmuch on palliative techniques (viz. what might help teachers feel better despite the pressures) rather than give enough weight to actions which teachers could take in order to better equip themselves with the skills required to deal with the demands on them more effectively (and hence the need for certain types of school-based in-service work and better working conditions and arrangements). The elaboration of organisational resources includes a very interesting section on improving staff selection, but there is a dilemma here which could have usefully been discussed· to what extent should the likelihood that someone might find their teaching post stressful mitigate against their appointment? Improving staff selection is presumably more to do with making sure those appointed can do their job well, rather than to make it harder for 'stress-prone' candidates (if they could be identified) to enter the profession. A number of particularly useful points are also made in discussing staff induction and development. Whilst Dunham's exploration of stress is very much concerned with in-service implications, there are also important links which the reader can make concerning initial teacher training.

This book is not the sort of book that slots primarily into the research literature on occupational stress generally; rather, it is aimed at teachers and other educationists interested in teacher stress from practical concerns about the quality of professional development and practice, and I am delighted that the author biased the book in this latter direction. For those interested in the former, there is however still much of interest.

In recent years there has been a real need to consider teacher stress in terms of a more general consideration of the pressures and developments facing both senior management and classroom teachers; this book addresses itself to such a consideration, and I highly recommend it.

CHRIS KYRIACOU

ENTWISTLE, N. J., and RAMSDEN, P. (1983). *Understanding Student Learning*. London: Croom Helm, pp. 248, £14·95.

How do successful students learn? In what ways do they differ from other students? What can be done to help the latter to be more like the former, and what can teachers do to help their students to become more effective learners? The apparent simplicity of these questions belies the enormity of the task of trying to find answers to them. In *Understanding Student Learning* Noel Entwistle and Paul Ramsden make a courageous attempt to do so in the context of higher education.

It is an exciting book. The publishers should be ashamed of the appalling standard of book-production, their worst sin being to place the lines (of unjustified typewriter script) so close together that the act of reading is thoroughly unpleasant. All the same, the struggle is worth it. We are told that the book contains the findings of the largest programme of research into student learning ever carried out in Britain. Entwistle and Ramsden have succeeded in their aim to describe the investigations in a form that is accessible to students, lecturers, and all who have an interest in education at degree level. They have given us a fascinating account of the widely varying ways in which students approach their courses.

A major conclusion is that of the various dimensions along which students' approaches differ, perhaps the most crucial is the *depth* at which learning takes place. Strictly speaking this is not a single dimension: they see a deep approach as being directed to the content of the material, and reflecting the particular knowledge, experience and interests of the individual learner. A surface approach, in contrast, leans towards the specific task and its requirements: it is characterised by situations in which the learner tries to impress alien material on memory for a limited period and with the specific intention of satisfying external demands. The findings show that to some extent these contrasting approaches can be seen as reflecting enduring styles of learning, but the authors admit that the demands of a particular task, as well as current motives, incentives and interests, and a student's existing expertise, as well as anxieties and expectancies about examinations, can all be important. And it is clear that there are many ways in which teachers can aid students by encouraging deeper approaches to learning. (Perhaps the most effective of all a teacher's actions are ones that help to make a learner less dependent upon teaching.)

Entwistle and Ramsden are conscious that at present there seems to be no one best route to understanding the causes and implications of students' differing approaches. They use a mixture of research methods, including interviews, questionnaires and inventories, occupying different positions in the quantitative-qualitative dimension. One source of possible confusion to the reader is the fact that different researchers have adopted alternative but overlapping terminologies. The survey of findings reveals identical terms used with very different meanings, and diverse terms representing apparently similar processes or approaches. To my mind some investigators in this area are too eager to leap in with factor-analytic techniques, not always with any strikingly useful outcome.

It will never be possible to give a simple prescription for successful approaches to learning in higher education. What works best depends too much upon particular circumstances and subject matter, the characteristics of an individual student, and the relationship between all of these. But sensitive investigations are revealing some important general principles and powerful generalisations this book is an invaluable guide to recent developments.

MICHAEL HOWE

FRUDE, N., and GAULT, H. (Eds.) (1984). *Disruptive Behaviour in Schools.* Bognor Regis: Wiley, pp. 234, £16·95.

LAWRENCE, J., STEED, D., and YOUNG, P. (1984). *Disruptive Children — Disruptive Schools?* London: Croom Helm, pp. 263, £14·95.

Neither book adequately comes to grips with providing an understanding of disruption within the enfolding contexts of schooling and society. There is then a tendency to focus upon the child who is disruptive or the acts of the child who is disruptive and how this might be cured or how the child might be readjusted in some way. There is little sense of turning the attention to the nature of schooling and reforming that in the interests of the child, the real interests of the child, not those interests defined by some adult as being "in the child's best interests". In that sense both books may be called unproblematically pragmatic and offer a number of practical hints and suggestions for teachers at various levels on what to do with the child who disrupts the control of the teacher or offer classificatory analyses of the disruptive incident.

Now, taking each separately, the Frude and Gault book offers quite a range of perspectives on disruption and in that sense may act as a useful source book. For example, the chapter by the Newsons offers a broadly statistical approach skilfully illuminated by parents' comments. The chapter by Woods offers an interactionist approach and focuses upon the questions "Who defines acts as disruptive? Which pupils commit disruptive acts? What is disrupted?" In this sense, then, the book offers the student a useful introduction to work in the broad area of disruption/deviance and, indeed, many of the articles are usefully scholarly.

The Lawrence, Steed and Young book provides a more detailed account of two research projects in 'disruption'. And disruption is defined as "behaviour out of place; the conditions for identifying it are a set of ordered relations and a contravention of that order. Disruption is a by-product of systematic ordering and classification, involving rejecting inappropriate elements" (p. 17). Unfortunately, there is little evidence on the part of the authors of a systematic critique of the order through which incidents and persons become defined as disruptive and through which certain persons come to have the power to make such definitions. If this order proves to provide a desirable environment for education from the points of view of all involved then perhaps one could find merit in developing a social technology of disruption control/prevention. If not, then the political implications of defining disruption from school-defined order make a technology of control and prevention unacceptable

JOHN SCHOSTAK

JEEVES, M. A., and GREER, G. B. (1983). *Analysis of Structural Learning.* London: Academic Press, pp. ix + 269, n.p.

The mathematician who opens this book will be excited to find that, according to the foreword by Jerome Bruner, it contains a "carefully crafted and bold investigation of the nature and origins of abstract thinking — the kinds of abstract thinking that are the base of logic and mathematics".

Two kinds of experiment are discussed Subjects operate equipment with coloured lights in an attempt to catch on to the "rule" by which the equipment has been programmed The possible rules in question are not, of course, revealed to the unfortunate subjects but we, the readers of the book, deserve a clearer explanation than that provided by the authors.

The first type of equipment concerns the possible rules for creating new sets which can be created from two sets A, B and their complements A′, B′. Since the universal set is made up of 4 disjoint sets AnB, AnB′, A′nB, A′nB′ and since each of these can be either used or not used in the rule, there are $2^4 = 16$ possible rules. These will be familiar to many readers as the 16 distinct truth tables which can be formed from two propositions P and Q. The subject has to guess, and describe, the "rule" being operated by performing repeated trials.

The second type of equipment is concerned with a single, but more complicated, "rule" about four symbols. Any two are combined (or "multiplied") to give a third. The rule chosen is in fact the same which governs the four numbers 1, 3, 5, 7 when multiplied modulo 8 (for example $3 \times 5 \equiv 7$ because $15 = 7 + 8$) but for the subjects the four symbols are unhelpfully called yellow circle, red circle, yellow triangle, red triangle. It is not surprising that the subjects (including an Older Adults group composed of 20 lecturers and professors) found some difficulty in discovering the rule for combination of symbols.

The results of both types of experiment are documented with a degree of detail normally associated with a research article rather than a book, and are surely above criticism. But the more important question is: what has all this to do with logic and mathematics? I am not convinced that there is much connection between understanding mathematics and catching on to mysterious "rules" by repeated trials, despite the authors' finding that there is a correlation between performance at their strange tasks and performance in mathematics examinations. It is hard to see much in common between current work in the psychology of learning mathematics, with its emphasis on understanding, and the authors' painstaking researches.

But surprisingly there is a link at a deeper level. A recurring theme in the psychology of learning mathematics is the sensitivity of questions to the way in which they are formulated Apparently minor changes in language, or in an accompanying diagram, can cause an enormous difference in success rates. Analogously, one of the main results in the present book is that the "rules" which were more quickly discovered by subjects were those which could be most simply described, given the limited language at their disposal. The conclusion towards which this surely points is that the correct aim of mathematics learning is not to guess mysterious "rules" but to acquire the language with which rules of mathematical structures become easy to express and therefore also easy to discover and to understand.

RALPH SCHWARZENBERGER

LEVY, P., and GOLDSTEIN, H. (Eds.) (1984). *Tests in Education: A Book of Critical Reviews.* London: Academic Press, pp. xxvii + 718, £20·00.

It was only recently that I discovered that Buros (pronounced *Bewross* or *Booross* in the U.K.) is *Burrows* in the U S.A. That gratuitous piece of intelligence sprang to mind when reviewing this edited volume, for the *Mental Measurement Yearbook* has long been the mental measurer's Bible. This collection of independent test reviews by a team of experienced reviewers from British higher education establishments is, as the editors explain, an attempt to make good the paucity of British educational tests in its illustrious predecessor. In view of the widespread use of tests, many with suspect or unspecified psychometric properties, the intention cannot but be applauded, and Levy and Goldstein are well-suited to the job

The authors' expressed intention was to focus upon "those tests which are more theoretically based or more technically developed". Others were excluded, not only because of dubious psychometric qualities but because they were "truly class tests, designed for informal use by teachers" but these, they aver, should not be ignored by the class teacher. The user must therefore determine whether the absence of a favourite test denotes its unreliable characteristics, its "true-class-test" qualities, or the consequence of a capricious postal service during compilation

Each test review is preceded by such information as is available on date of origin, copyright dates for test and manual, date of standardisation, and the dates on which revisions were undertaken. Then follows the review, itemised according to the contents of the test, the purposes for which the test is intended, item preparation, administration, standardisation, sex differences, reliability, validity, score interpretation and test use. The editors' comments on some of these issues are a revealing 'state of the art' commentary and an important section of the book. They should not be ignored as an unnecessary gloss upon the 'useful' content. The Introduction would make useful preliminary reading for any teacher or student about to embark on a study of tests and testing. It has an authoritative but down-to-earth quality which appeals.

In describing in an admirably jargon-free way the uses to which tests may be put, the editors comment that of the tests "very few have any clearly useful statement of educational rationale", and that so-called 'diagnostic' tests sometimes trust that an "experienced teacher" will be able to "make some sense of the signs"

On the problem of administering tests, and the resultant effects upon reliability of variations in administration procedures, the authors note that there is evidence of frequent lapses amongst testers. "The less experienced test user . . . needs to be told clearly what to do . . ." They do not refer to what I suspect to be an equivalent problem amongst "old hands" — familiarity breeding abuse. Am I alone in being dismayed at the variations that have sneaked in to a WISC protocol I thought I knew perfectly?

The collection of reviews covers early development, language, mathematics, composite attainments, general abilities, personality and counselling, and "other topics" (encompassing, amongst others, science reasoning, art and music tests). Some reviews are brief, but most are extensive and seem to be thorough. In two cases tests have been reviewed by two independent contributors and in some vague hope of a crude estimate of consensus, or maybe a malicious hope of contradiction, one pair were given special scrutiny.

Young's Group Reading Test was reviewed by Caroline Gipps and Christine Mabey. General descriptions of the test, its contents and purpose are almost identical, but under "item preparation" Gipps was somewhat more overtly critical than Mabey. She argued that no information was given on the item sampling frame and that reference to the elimination of strong sex bias was "vague in the extreme". Both agreed that the administration instructions were thorough, but on the issue of standardisation Gipps was again rather more critical Both reviewers found scoring straightforward, although Mabey gave more information, as she did on reliability, validity, and interpretation. The agreement between reviewers also extended to the general evaluation. Gipps was perhaps a little less enthusiastic than Mabey, implying that the cheapness and ease of administration made it popular with LEAs, and suggesting that it was useful as a "rough guide to overall standards ... with a caveat about standardisation". She questioned the extensive use of synonyms in a reading test and opined that the test provided "little feedback for the teacher". Mabey was somewhat more approving in tone, if not in content. She declared it to be "rightly popular, not only on grounds of expense and ease of administration, but because of clarity in instructions and attractiveness to children". She also cautioned on the inadequacy of standardisation data and on the 'ceiling effect' possible when used with older children.

So the two reviewers seemed to agree very largely, although I felt that the tone of one was less enthusiastic than that of the other. However, there are many factors to be considered before a test is actually used in the classroom, and the tone of a review may pale into insignificance beside them. It is to be hoped that the content of the reviews will not, for within this book is a wealth of valuable information I cannot imagine that many education libraries, resource centres, teachers' centres, or psychological services will be without it.

GEOFFREY BROWN

PERERA, K. (1984). *Children's Writing and Reading: Analysing Classroom Language.* Oxford: Basil Blackwell, pp. 354, £8 95.

Perera claims that her book should provide teachers (and others such as speech therapists and educational psychologists) with the framework of knowledge they need in order, first, to assess their pupils' grammatical abilities and, secondly, to intervene appropriately to extend them (as advocated by the Bullock Committee).

Chapter Two provides a simplified grammar — based on Quirk *et al* and Quirk and Greenbaum — to give teachers a working vocabulary for analysis of children's oral and written language. This section, like the rest of the book, is well-organised, easy to read and carefully organised through the use of examples. Such a description of grammar, as well as being essential to the development of the book, may be of particular importance since many recently trained teachers will not have studied grammar; they will be lacking in the explicit, systematic foundation of grammatical knowledge which Perera suggests is necessary to identify and understand both strengths and weaknesses in their pupils' developing mastery of language. They will consequently be unable to monitor children's language development in a worthwhile manner.

The grammatical framework is used as a basis for the description in Chapter Three of children's acquisition of grammar in the mother tongue from the ages of 18 months to about 14 years. The particular focus is on those aspects of grammar that are still developing when a child starts school and on structures that cause children to struggle in their comprehension or production of speech

Perera points out (in Chapter Four) that oral and written language are very different: whereas this is neither original nor startling, what is useful is the clarity with which she demonstrates both the difference and the significance attached to it. This clarity is a direct consequence of the earlier grammatical description and indicates the importance of educators being able to identify and refer to these aspects; and thereby being aware of how the acquisition of grammar in oral and written modes interrelate and reinforce each other. Perera's analysis of this mutual dependence in the context of a developmental model provides a justification for, for example, reading aloud to all pupils and for teachers being sensitive when marking written work to the uncertain progress children make through the stages of grammatical development.

In so far as such conclusions seem fairly mundane, this book would not provide much satisfaction to those who scan it for 'tips for teachers'; it is, rather, in the opportunities it provides for teachers to analyse and discuss children's work at a non-trivial level that its importance is to be found. But the difficulty here is that, despite being easy to read, the book is not easy to digest as a whole. Equally, teachers who are attracted by the title could not be blamed for giving up before the questions of writing and reading are specifically addressed in Chapters Five and Six. Furthermore, since much of the book comprises an extensive summary of research which is too detailed for the teacher, yet not sufficiently rigorous for the researcher, it may not be fully satisfactory to any audience However, it deserves careful analysis and interpretation by those concerned with the training of teachers — particularly so at a time when the publication of English from 5-16 has expressed the uncertainties of contemporary discussion about the aims and methods of teaching English. The uncertainty persists in Perera's contribution, for it is still not clear whether teachers should, for instance, intervene in order to help 8-year-olds perform like 10-year-olds, but, in attempting to describe the processes children use in composing, it provides a starting point for working out a rule system and a theory strong enough to be tested

LIZ ROTH

QUIGLEY, S. P , and PAUL, P V. (1984). *Language and Deafness*. London: Croom Helm, £19·95.

This is an extremely well-documented and referenced text which should prove extremely valuable to all those specially interested in the problems hearing-impaired children have in learning language. It should be of special interest to students and researchers in the field of hearing handicap. Less serious readers may however, feel that the heavy referencing leads to a text which is somewhat fragmented and occasionally heavy going. Each chapter provides a useful summary and references are together at the end of the text.

The opening chapter outlines historical perspectives and provides clear and concise definitions, ensuring that the reader has a clear understanding as to the type of hearing-impaired children coming under discussion. They are children with losses in excess of 90 dB, deafness so great, suggest the authors, that even with good amplification, vision becomes a main medium of communication. An interesting discussion follows on various communication systems and an excellent chapter is rounded off by a list of principles representing what the author sees as the best practice in relation to language development in current educational programmes

Chapter 2 on Cognition and Language examines the major areas. Firstly, whether qualitative or quantitative differences exist between hearing-impaired and hearing subjects and, secondly, whether hearing-impaired subjects can provide us with knowledge about the relationship between language and thought.

In a discussion on primary language development, the authors trace the development of both normally hearing and hearing-impaired children. The description is done in relation to the medium of communication through which it is developed and includes Oral English, Manually Coded English, and American Sign Language. The discussion is even-handed and indicates across the media the delayed, but similarly sequenced language acquisition. The authors note and caution readers with regard to the definitions of approach. They argue that under each of the headings there is a wide range of practice: the oral programmes that are not entirely oral, the manually coded English programmes which do not adhere strictly to English structure in the manual component of the system. The authors conclude this section by suggesting that a pressing need for researchers is to establish whether or not certain skills can be transferred from a signed language (e.g., American Sign Language), to a spoken language such as English. This, they argue, is vital if the ability to read and write English is a desirable goal in the education of deaf children.

Chapters 4 and 5 on reading and writing point to the apparent immunity of these aspects of language to different approaches to encourage their growth. They argue that, in spite of increased understanding of many of the processes involved, hearing-impaired children (in general) have not shown improved performance in these skills. They also suggest the possibility of using other means of transmission of the information contained in the printed word, e.g., signed videotapes in American Sign Language.

There is an interesting chapter on bilingualism with English being taught as a second language. The authors see such an approach as valid, providing it is properly initiated and carefully evaluated. Barry Jones provides a chapter on language assessment and instruction drawing mainly on American material and the book is rounded off with an excellent chapter of conclusion.

This is a most important book and I am sure it will be widely read by all those who have an interest in deafness. The authors are noted for their balanced views and this text again reflects that fact, with demands for approaches to be researched and evaluated, and tailored to the widely differing needs of hearing-impaired children. There is no single approach capable of satisfying the needs of the group at which this book is aimed, those with the severest levels of sensori-neural deafness

IVAN TUCKER

STEPHENSON, G. M., and DAVIES, J. H. (Eds.) (1984). *Progress in Applied Social Psychology, Vol. 2.* Chichester: Wiley, pp. 325, £32·95.

The second in Stephenson and Davies' series has ten chapters, written by an international collection of authors. The editors group the chapters into three sets: "Institutions, professions and organisations", "Group decision-making" and "Psychology and the law".

In the first, rather loosely-bounded section we have a pessimistic review of the psychology of unemployment by Hartley and Fryer They point out that psychologists have been too much concerned with the effects rather than the causes of unemployment, and that received wisdom about the nastiness of unemployment is in fact based on a not-particularly large number of reliable studies. Hartley and Fryer's review is followed by a chapter with a very different feel, as two authors (Kiesler and Sibulkin) present a technical report (on a study of psychiatric admissions) which was, to this reviewer, uninvitingly dense. The third chapter is rather a curiosity, reporting on professional and academic psychologists' perceptions of each other in Italy. The occasionally fractured English seems to have been left uncorrected by the editors, with often rather charming results.

The fourth chapter, by W. P. Robinson, is the best written of the collection, which is fortunate for BJEP readers since it's about a topic close to their hearts. Robinson reviews, with some style, what we know (and don't know) about classrooms considered as places where someone has to be efficiently taught something. Robinson is not one to mince his words; if you suspect someone of being complacent about educational psychology, shove this chapter under his nose.

The last chapter in the section is on 'pluralistic ignorance'. It features some nice scholarly detective work and good updating of research, but, oddly, it fails to acknowledge any of the huge literature in attribution theory which bears on people's estimation of consensus.

The second section, group decision-making, has three chapters. The first, by Brehmer on judgment in small groups, is the chapter in this book that most nearly approximates what I suppose is my ideal of applied social psychology: substantial theory, with clear predictions, applied to real life. As it happens, unfortunately, it falls down on the last bit, reviewing, after exensive theoretical discussion, only a small handful of applied studies. The next chapter, by Komorita, is on justice and power, but leaves out even the handful of applied studies and replaces them with lab work. The last chapter in the section, by Rabbie and Oostrum, completes the slide by concentrating on two laboratory experiments, verbosely described.

The last section is graced by Sally Lloyd-Bostock's very readable prose. Her chapter, on psychologists and lawyers, is one of the most sensible of the book, and the message of qualified enthusiasm is a general one. The chapter could be read by anyone interested in pairing psychology up to some other discipline and mental lifestyle. The last chapter of the book, by Walker and Lind on the study of legal procedures, is a mixture of an empirical report (of lab work) and a 'theory' of procedure.

Overall, the reader new to applied social psychology may well feel that there are three types of article that make up the bulk of the literature. the pessimistic review; the atheoretical field study; and the optimistic laboratory experiment with limited external validity. This book is no more than a fair reflection of the truth.

CHARLES ANTAKI

TIZARD, B., and HUGHES, M. (1984). *Young Children Learning*. London: Fontana, pp. 286, £2·95.

This is certainly an interesting book, another 'shot in the locker' some will see it as. It provides concurrent validity to earlier work by Tizard and by Wood and chimes well with the current vogue for psychologists criticising nursery teachers. This is a readable book, well-organised and written with a certain degree of elegance and authority which I have come to expect from the principal author. The balance, too, between child conversation and subsequent analysis is attractively sustained, leading one on easily to related levels of the discussion and contemporary research.

But it is an account with which I feel uneasy and which I fear could have a mischievous effect. What then is this reviewer's problem? After all, twelve years as a colleague of Gordon Wells should have made me familiar with arguments concerning the relatively high quality of almost all mother-child conversation regardless of social class, a point made abundantly clear in the book. Moreover, the general myth of *oral* language deficit, rather than difference in the value/attitudinal context in which the language is embedded, does need repeated denial; and that denial comes systematically and clearly throughout chapter six. Perhaps, then, my unease is generated 'between the lines'; perhaps it is partly to do with Whitehead's comment to the effect that if research does not quite fit one's own perceptions of reality, then treat that research with caution; perhaps it is to do with a hint that the authors do not sufficiently empathise with (as distinct from understand) the problems of 'day-to-day system maintenance' involved in teaching.

The research resulting in this volume is based upon data derived from observations of 30 girls, 15 from two distinct social class groups; all of them aged 3 years 9 months, all of them with two adult caretakers or parents; all of them speaking English as their mother tongue. It really ought therefore, to be called, *Young Girls Learning*. Conversations at home and at school were recorded and analysed Tables are presented in the appendix. The whole refers to, and falls in line with previous research by Tizard, Hughes, *et al*. Consequently, and as I have indicated, the book carries authority based on research experience.

However, even with all that experience, I am not sure I would sweepingly categorise the British nursery school as having days 'largely given over to *free* play' [p. 180 — my italics]; nor would I care to sound such a clear message that the nursery school is deficient, when compared with the home, in respect of fostering intellectual development. But the authors have no such qualms. They present a view, well-supported by theirs and others' research, of teacher-child conversations being more limited and limiting than in the home. How will such evidence be interpreted and employed? Tizard and Hughes rightly point out that there are unsatisfactory gaps between the children's learning at home and that at school. They say, too, that there is little British evidence to suggest that nursery school does stimulate language and intellectual development in the way many have believed and asserted. They take the obligatory side-swipe at Piaget which 'card-carrying psychologists' make nowadays to establish their credentials.

Like Bennett's 1976 book, this volume has received a fair degree of advance publicity. I hope, however, the data in this book are not as simplistically interpreted. If these data are seen as showing *some* of the problems, if politicians don't interpret 'widening children's horizons' and extending their general knowledge (p. 261) as a signal to either increase didactic schooling or close nurseries altogether, then it will probably do some good in the long run. But, I remain uneasy Assessment? A for effort, D for empathy!

PHILIP GAMMAGE

PUBLICATIONS RECEIVED

Books for review have been selected from the following
('n.p.' means 'no price mentioned').

ALTMAN, I., and CHEMERS, M. M., (1985). *Culture and Environment.* Cambridge: Cambridge University Press, pp. 337, £6·95.

AMIR, Y., and SHARAN, S. (1984). *School Desegregation.* London: Lawrence Erlbaum Associates, pp. 267, £23·00.

APPLEBEE, A. N. (1984). *Contexts for Learning to Write: Studies of Secondary School Instruction* London: Ablex, pp. 224, £29·15.

BOLT, R. (1985). *Penguin Passnotes. A Man for All Seasons.* Harmondsworth: Penguin Books, pp. 96, £0·95.

BOLT, S. (1985). *Penguin Masterstudies Hamlet.* Harmondsworth: Penguin Books, pp. 96, £1·75.

BRENNAN, W. K. (1985). *Curriculum for Special Needs.* Milton Keynes: Open University Press, pp. 213, p. £5·95, cased £15·00.

CONNOR, P. J. (1985). *Penguin Passnotes. Cider with Rosie.* Harmondsworth: Penguin Books, pp. 79, £0 95

COOTE, S. (1985). *Penguin Masterstudies: The Waste Land.* Harmondsworth: Penguin Books, pp. 142, £1·75.

COOTE, S. (1985) *Penguin Passnotes: The Prologue to the Canterbury Tales.* Harmondsworth: Penguin Books, pp. 166, £1·50.

COOTE, S. (1985). *Penguin Passnotes: As You Like It.* Harmondsworth: Penguin Books, pp. 94, £0·95.

CUNNINGHAM, J. (1985). *Penguin Masterstudies The Miller's Tale.* Harmondsworth: Penguin Books, pp. 118, £1·75.

DAVES, C. E. (Ed.) (1984). *The Uses and Misuses of Tests.* London: Jossey-Bass, pp. 133, n.p.

ENTWISTLE, N. (1985). *New Directions in Educational Psychology 1.* Basingstoke: Falmer Press, £16·95 (hb), £9·95 (pb).

EVANS, K. M. (1984). *Planning Small Scale Research.* Windsor: NFER-Nelson, pp. 64, £3·95

FAMILY RIGHTS GROUP (1984). *Permanent Substitute Families.* London: Family Rights Group, pp. 31, £2·00.

FISHBEIN, H. D. (1984). *The Psychology of Infancy and Childhood.* London: Lawrence Erlbaum Associates, pp. 435, £45 00.

FLEMING, J. (1985). *Blacks in College.* London: Jossey-Bass, pp 276, £20 95.

GHOLSON, J. B., and ROSENTHAL, T. L. (1984). *Applications of Cognitive-Developmental Theory.* London: Academic Press, pp. 295, n.p.

GOSLING, D., and MUSSCHENGA, B. (1985). *Science Education and Ethical Values: Introducing Ethics and Religion into the Science Classroom and Laboratory.* Switzerland: WCC Publications, pp. 115, £4·75.

HALL, W. S., NACY, W. E., and LINN, R. (1984). *Spoken Words Effects of Situation and Social Group on Oral Word Usage and Frequency.* London: Lawrence Erlbaum Associates, pp. 489, £37·95.

HANDLEY, G. (1985). *Penguin Masterstudies: Vanity Fair.* Harmondsworth: Penguin Books, pp. 157, £1 75.

HANDLEY, G. (1985). *Penguin Passnotes. To Kill a Mockingbird.* Harmondsworth: Penguin Books, pp. 105, £0·95.

HUBERMAN, A. M , and MILES, M. B. (1984). *Innovation Up Close: How School Improvement Works.* New York. Plenum, pp. 309, $35·40.

JAMIESON, G. H. (1984). *Communication and Persuasion.* Beckenham: Croom Helm, pp. 170, £14 95.

KENNER, C. (1985). *No Time for Women: Exploring Women's Health in the 1930s and Today.* London: Pandora Press, pp. 168, £4·95.

KIERAS, D. E. and JUST, M. A. (Eds.) (1984). *New Methods in Reading Comprehension Research.* London: Lawrence Erlbaum Associates, pp. 377, £22·95.

KOSHY, K T. (1985) *Revision Notes on Psychiatry.* Sevenoaks: Hodder & Stoughton, pp. 248, £4·50.

MARTON, T., HOUNSELL, D., and ENTWISTLE, N. (1985) *The Experience of Learning.* Edinburgh: Scottish Academic Press, pp 242, £9·75.

McLaughlin, B. (1984). *Second Language Acquisition in Childhood Vol. 1.* London: Lawrence Erlbaum Associates, pp. 261, £23·00.

Nicol, A. R. (1985). *Longitudinal Studies in Child Psychology and Psychiatry. Practical Lessons from Research Experience.* Bognor Regis: Wiley, pp. 411, £29·50.

Nixon, J. (1985). *A Teacher's Guide to Multicultural Education.* Oxford: Basil Blackwell, pp. 183, £5·95.

Nuttin, J. (1984). *Motivation, Planning and Action.* Leuven: Leuven University Press, pp. 251, 1,200 Belgian francs.

Oosthoek, H., and Van Den Eeden, P. (1984). *Education from the Multi-Level Perspective.* New York: Gordon and Breach, pp. 281, $39·95.

Owens, R. G., and Ashcroft, J. B. (1985). *Violence: A Guide for the Caring Professions.* Beckenham: Croom Helm, pp. 198, £14·95 (hb), £7·95 (pb).

Pollitt, A., Entwistle, N., Hutchinson, C., and De Luca, C. (1984). *What Makes Exam Questions Difficult? An Analysis of 'O' Grade Questions and Answers.* Edinburgh: Scottish Academic Press, pp. 185, £8·75.

Quicke, J (1984). *Disability in Modern Children's Fiction.* Beckenham: Croom Helm, pp. 176, £12·95.

Quilliam, S. (1985). *Penguin Passnotes. Romeo and Juliet.* Harmondsworth: Penguin Books, pp. 108, £0·95.

Reynolds, C R., and Brown, R. T. (1984). *Perspectives on Bias in Mental Testing.* New York: Plenum Press, pp. 608, $50·00

Schopler, E., and Mesibov, G. B. (1984). *The Effects of Autism on the Family.* New York: Plenum Press, pp. 384, $35·00.

Seltman, M., and Seltman, P. (1985). *Piaget's Logic; A Critique of Genetic Epistemology.* Hemel Hempstead: Allen and Unwin, pp. 353, £30·00.

Sharan, S. (1984). *Cooperative Learning in the Classroom.* London: Lawrence Erlbaum Associates, pp. 160, £12·50.

Sternberg, R J. (1985). *Beyond IQ a Triarchic Theory of Human Intelligence.* Cambridge: Cambridge University Press, £25 00 (hb), £8·95 (pb).

Taylor, N. (1985). *Penguin Passnotes. Julius Ceasar* Harmondsworth: Penguin Books, pp. 93, £0·95.

Vance, C. (1985). *Pleasure and Danger.* London: Routledge and Kegan Paul, pp. 462, £12·95 (cloth), £6 95 (paper).

Waldron, T. P. (1985). *Principles of Language and Mind An Evolutionary Theory of Meaning.* London: Routledge and Kegan Paul, pp. 232, £14·95.

Walford, G. (1984). *British Public Schools: Policy & Practice.* London. Taylor & Francis, pp. 220, £14·50 (cloth), £8·50 (paper)

Winston, R. B. *et al.* (1984). *Developmental Academic Advising.* London: Jossey-Bass, pp. 571, £33 50.

Zeisel, J (1985). *Inquiry by Design: Tools for Environment-Behaviour Research.* Cambridge Cambridge University Press, pp. 250, £6·95.

Zube, E. H. (1985). *Environmental Evaluation. Perception and Public Policy.* Cambridge: Cambridge University Press, pp. 148, £5·95.

Rhyme and Reason in Reading and Spelling

Lynette Bradley and Peter Bryant

Children have been learning nursery rhymes for centuries. But they are more than just child's play. The authors of this new book report on their four-year longitudinal study of 368 children. That study identifies a causal relationship between reading ability and a prior ability to categorize words on the basis of sound. A child's sensitivity to rhyme and alliteration will affect his later ability to learn to read and spell. The study also shows that one may identify and foretell a child's learning difficulties based on his ability to use rhyme. The authors report on intervention methods used with a test group of the children in their study and suggest ways in which teachers may use word games to help children overcome their learning problems. This is the first volume in the International Academy for Research in Learning Disabilities Monograph series.

In U.S.A. $7.95

Customers in the United Kingdom and Eire should place orders and obtain price information from Transatlantic Book Service (TABS), 24 Red Lion Street, London WC1A 4PX, England.

The University of Michigan Press
Ann Arbor, Michigan, U.S.A.

Journal of Occupational Psychology

Special issue on 'Psychology and Industrial Relations'

Papers are invited for a special issue of the Journal, to be published in late 1986. Contributions are welcomed on any psychological aspect of industrial relations, in particular on psychology and trade unionism, intergroup relations, commitment and mobilization, public opinion, discrimination, and the role of law. Review articles will also be welcomed.

The special issue will be edited by joint guest editors, Dr John Kelly (London School of Economics) and Dr Jean Hartley (Warwick University).

Four copies of submissions should be sent to John Kelly at the London School of Economics, Houghton Street, London WC2A 2AE.

Submissions should arrive no later than 31 October 1985.

The *Journal of Occupational Psychology* is edited by Dr David Guest. ISSN 0305-8107 Volume 58 (1985), price £41.00 (US$ 79 50) is available from:

The British Psychological Society
The Distribution Centre, Blackhorse Road, Letchworth, Herts SG6 1HN, UK

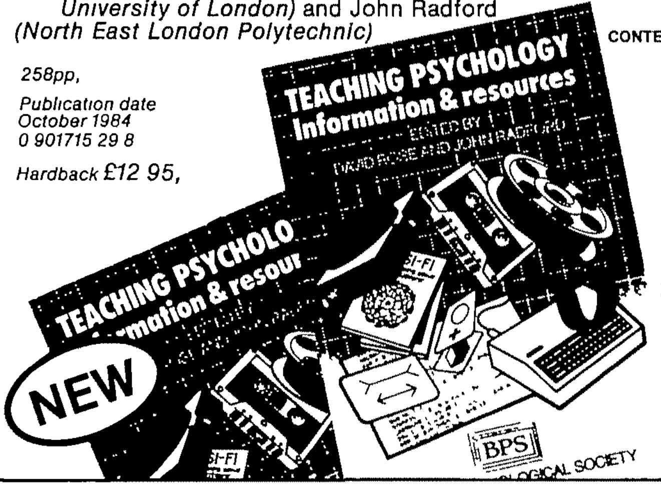
TEACHING PSYCHOLOGY: INFORMATION AND RESOURCES

Edited by David Rose *(Goldsmiths' College, University of London)* and John Radford *(North East London Polytechnic)*

258pp,

Publication date October 1984
0 901715 29 8

Hardback £12 95,



Designed for teachers of psychology in secondary, tertiary and higher education, this unique book is a compendium of sources of information and equipment, as well as a guide to innovative teaching strategies The editors, acknowledged experts in the field of psychology teaching, have brought together contributors from both academic and technical departments in colleges and universities. This is the first publication of its kind outside the USA

CONTENTS· *Libraries, publishers and specialist bookshops ● Audio-visual equipment (incl suppliers) ● Laboratory equipment (incl suppliers) ● Do-it-yourself laboratory equpment ● Microcomputers in psychology teaching ● Psychological tests ● Feature films in psychology teaching (incl distributors) ● Literary resources ● Approaches to group-dynamics ● Bibliography of the literature on teaching psychology in the UK and USA ● Careers (incl useful addresses) ● Ethical considerations*

READERSHIP:
Teachers up to undergraduate level, Including colleges of FE, institutes of education and adult education centres, Teaching-hospital staff, Management trainers, Personnel officers, Advisory teachers, Information and resources officers, School and university librarians

THE BRITISH PSYCHOLOGICAL SOCIETY

All orders to The Distribution Centre, Blackhorse Road, Letchworth, Herts SG6 1HN, UK

PUBLICATIONS OF THE BRITISH PSYCHOLOGICAL SOCIETY

British Journal of Psychology
British Journal of Medical Psychology
Journal of Occupational Psychology
British Journal of Social Psychology
British Journal of Clinical Psychology
British Journal of Developmental Psychology
British Journal of Mathematical and Statistical Psychology
The British Journal of Educational Psychology

The *British Journal of Psychology*, the *British Journal of Medical Psychology*, *Journal of Occupational Psychology*, the *British Journal of Developmental Psychology*, the *British Journal of Social Psychology* and the *British Journal of Clinical Psychology* appear in quarterly parts. The 1985 subscription prices, payable in advance, are £60 00 (US$110 00) for the *British Journal of Psychology*; £38·00 (US$64·50) for the *British Journal of Social Psychology*, £38 00 (US$64 50) for the *British Journal of Clinical Psychology*, £35 00 (US$64 50) for the *British Journal of Developmental Psychology*, £49 00 (US$90 50) for the *British Journal of Medical Psychology*; and £41 00 (US$79 50) for the *Journal of Occupational Psychology*

The *British Journal of Mathematical and Statistical Psychology* is issued twice a year, in May and November The subscription rate, payable in advance, is £41·00 per volume post free (US$79·50 overseas).

Subscriptions to any of these journals may be sent to a bookseller or subscription agent, or to The British Psychological Society, The Distribution Centre, Blackhorse Road, Letchworth, Herts SG6 1HN, UK. A special price of £273 00 (US$500 00) for all seven journals going to the same address is offered for 1985.

The British Journal of Educational Psychology is issued three times a year, in February, June and November. The prepaid subscription is £18·00 a year, post free, or £6 50 each number. The prepaid subscription price in the USA is $40 00 post free Subscriptions should be sent to the Scottish Academic Press, 33 Montgomery Street, Edinburgh EH7 5JX

MANUSCRIPTS

For the *British Journal of Psychology* should be sent to Professor D E. Blackman, Department of Psychology, University College Cardiff, PO Box 78, Cardiff CF1 1XL

For the *British Journal of Medical Psychology* should be sent to Dr Sidney Crown, Department of Psychiatry, The London Hospital, Whitechapel, E1 1BB

For the *Journal of Occupational Psychology* should be sent to Dr D. Guest, Department of Industrial Relations, London School of Economics and Political Science, Houghton Street, London WC2A 2AE

For the *British Journal of Social Psychology* should be sent to Dr A. S. R Manstead, Department of Psychology, The University, Manchester M13 9PL.

For the *British Journal of Clinical Psychology* should be sent to Dr D Shapiro, MRC/SSRC Social and Applied Psychology Unit, Department of Psychology, The University, Sheffield S10 2TN

For the *British Journal of Mathematical and Statistical Psychology* should be sent to Dr C. Leach, Department of Psychology, University of Newcastle upon Tyne, Newcastle upon Tyne NE1 7RU.

For the *British Journal of Developmental Psychology* should be sent to Professor P E. Bryant, Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford OX1 3UD.

For *The British Journal of Educational Psychology* should be sent in duplicate to Professor Geoffrey Brown, School of Education, University of East Anglia, Norwich NR4 7TJ

Contributors are advised to consult *Suggestions to Contributors* issued by the Standing Committee on Publications of the British Psychological Society and obtainable, price £2·00 post free, from The British Psychological Society, St. Andrew's House, 48 Princess Road East, Leicester LE1 7DR.

All manuscripts should be submitted in duplicate.

OFFPRINTS

Contributors on application receive fifty copies of their papers free Additional copies may be obtained in units of fifty copies.

REPRINTS

The *Journal* is available in reprint from Swets and Zeitlinger B V, P.O. Box 810, 2160 SZ Lisse, The Netherlands, Vols 21-29, 31-47

SUBSCRIPTIONS TO THE BRITISH PSYCHOLOGICAL SOCIETY

The subscription for Fellows, Associates and Graduate Members is £30 00

They are entitled to purchase any or all of the Society's Journals at the special rates of £6 00 per Journal per annum.

Information on membership and activities may be obtained from the Secretary, The British Psychological Society, St. Andrew's House, 48 Princess Road East, Leicester LE1 7DR.

BRITISH JOURNAL OF EDUCATIONAL PSYCHOLOGY,

Vol. 55, Part 2, 1985

CONTENTS

Articles:

Levy-Shiff, R., and Hoffman, M. A. — Social Behaviour of Hearing-Impaired and Normally-Hearing Preschoolers . . . 111
Woodhouse, D. A., Hall, E., and Wooster, A. D. — Taking Control of Stress in Teaching . . . 119
Entwistle, N. J., and Kozéki, B. — Relationships between School Motivation, Approaches to Studying, and Attainment among British and Hungarian Adolescents. . . . 124
Craske, M. L. — Improving Persistence through Observational Learning and Attribution Retraining . . . 138
Kyriacou, C. — Conceptualising Research on Effective Teaching . . . 148
Mevarech, Z. R. — The Relationships between Temperament Characteristics, Intelligence, Task-Engagement and Mathematics Achievement . . . 156

Research Notes:

Kember, P. — The Transcendental Meditation Technique and Postgraduate Academic Performance . . . 164
Hart, J. G. — LAWSEQ: Its Relation to Other Measures of Self-Esteem and Academic Ability . . . 167
Book Reviews . . . 170
Publications Received . . . 183

26 MAY 1986

CALCUTTA UNIVERSITY

UK ISSN 0007-0998

THE BRITISH JOURNAL OF EDUCATIONAL PSYCHOLOGY

NOVEMBER 1985

VOL. 55, PART 3

CALCUTTA UNIVERSITY
CENTRAL LIBRARY

THE
BRITISH JOURNAL
OF
EDUCATIONAL PSYCHOLOGY

Edited by
Professor GEOFFREY BROWN

Assistant Editors:
Dr. R. L. BURDEN
Dr. C. W. DESFORGES
Dr. D. FONTANA

Sub-Editor:
Dr. S. CUNNINGHAM

Book Review Editor:
Dr. C. W. DESFORGES

Advised by the Editorial Committee:
Dr. H. BELOFF
Professor M. B. SUTHERLAND
Professor J. WRIGLEY

Vol. 55, Part 3

The British Journal of Educational Psychology is issued by the British Psychological Society once a term, in February, June and November. The subscription is £18 00 a year post free, payable in advance, or £6 50 each number.

Subscriptions and Orders should be sent to Scottish Academic Press, 33 Montgomery Street, Edinburgh, EH7 5JX Members of the British Psychological Society receive the *Journal* on special terms.

Papers for publication in duplicate should be sent to Professor Hazel Francis, University of London Institute of Education, 20 Bedford Way, London WC1H 0AA Guidance to contributors on the preparation of papers for the *Journal* is given at the end of the November number each year.

All *Books for Review* for *The British Journal of Educational Psychology* should be sent to Dr C. W Desforges, St. Martin's College, Lancaster LA1 3JD

Advertisements for this *Journal* should be sent to Scottish Academic Press.

Published by
SCOTTISH ACADEMIC PRESS, 33 MONTGOMERY STREET, EDINBURGH, EH7 5JX,
for the British Journal of Educational Psychology, Ltd

All rights reserved

© 1985. The British Journal of Educational Psychology

PRINTED IN GREAT BRITAIN BY CLARK CONSTABLE, EDINBURGH AND LONDON

Br. J. educ. Psychol., 55, 185-212, 1985

THE ROLE OF METALEARNING IN STUDY PROCESSES

By J. B. BIGGS
(*Department of Education, University of Newcastle, Australia*)

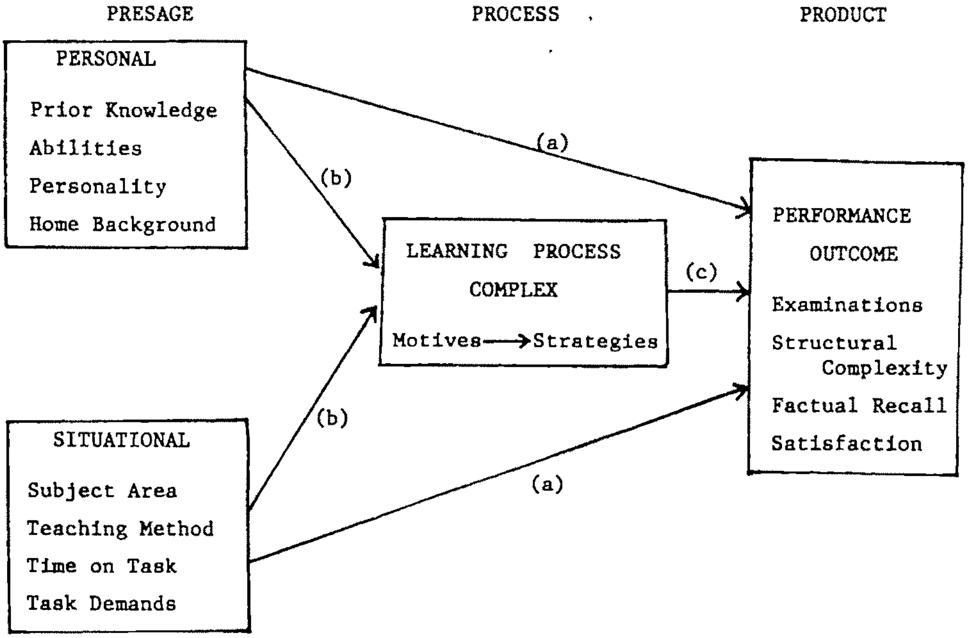
SUMMARY. Effective learning under institutional conditions requires, first, that students are aware of task demands and of their intentions of how, or even whether, to meet those demands, and, second, that they assess realistically, and exert control over, their own cognitive resources. The fulfilment of such conditions involves a sophisticated kind of metacognition, here called *metalearning*. The present paper describes a series of studies that collectively explicate the development and role of metalearning in the learning and study processes of secondary and tertiary students. Ability patterns, locus of control, variety and quality of certain non-school experiences, and extent and kind of motivation all seem to be involved in the development of metalearning capability. A model of student learning is then described, in which personal and situational factors are linked to performance by three main approaches to learning: deep, achieving, and surface. These approaches involve varying degrees of metalearning and lead to qualitatively different learning outcomes. An intervention study on improving student learning is described that tests important aspects of the model, and further implications for teaching and research are suggested.

INTRODUCTION

IN an earlier paper, the present writer described a three-stage model of student learning (Biggs, 1978), the essentials of which are outlined in Figure 1.

FIGURE 1

GENERAL MODEL OF STUDENT LEARNING



Presage factors are of two kinds. Personological factors comprise enduring personal characteristics such as abilities, cognitive styles, and previous experience; situational ones comprise course structures, evaluation procedures, time on task, and the like.

Process factors make up the "learning process complex", which comprises three basic approaches to learning, each approach consisting of a motive and a related congruent strategy.

The *product,* performance, is determined by the presage factors, personal and situational, both directly (indicated by the (a) arrows), and indirectly as mediated by the learning process complex (indicated by the (b) and (c) arrows). This paper is concerned with these indirect effects, referring as they do to metacognitive processes, and in particular to students' perceptions of their own cognitive resources, motivations, and the task demands.

This basic model has been elaborated (Biggs, in press; and as discussed below), and the approaches renamed to bring them into line with usage by other writers. The learning process complex is the focus of the model, and is constituted as outlined in Table 1.

TABLE 1

MOTIVE AND STRATEGY IN APPROACHES TO LEARNING AND STUDYING

Approach	Motive	Strategy
SA: *Surface*	Surface motive (SM) is instrumental: main purpose is to gain a qualification with pass-only aspirations, and a corresponding fear of failure.	Surface strategy (SS) is reproductive: limit target to bare essentials and reproduce through rote learning
DA· *Deep*	Deep motive (DM) is intrinsic: study to actualise interest and competence in particular academic subjects.	Deep strategy (DS) is meaningful· read widely, inter-relate with previous relevant knowledge
AA: *Achieving*	Achieving motive (AM) is based on competition and ego-enhancement. obtain highest grades, whether or not material is interesting.	Achieving strategy (AS) is based on organising: follow up all suggested readings, schedule time, behave as "model student".

The following points should be noted:

(1) A formal learning situation generates three common expectations, which in turn shape the student's motives for engaging the task: to obtain a qualification with minimal effort, to actualise one's interests, and to manifest one's excellence publicly by obtaining the highest grades. These motives in turn are usually associated with cognate strategies: to reproduce what is perceived to be essential data, to understand the meaning of the task, and to organise one's time in order to optimise available time on task, respectively. The motives and strategies collectively comprise three broad *approaches* to learning, which have been observed not only by the present writer but (with minor differences on the first in particular) independently by Entwistle *et al.* (1979), Watkins (1983), O'Neil and Child (1984) and Entwistle and Kozéki (1985). There is, then, considerable support for postulating these three approaches, but this is not to say that these are the only three. Entwistle and Watkins, in the above and in other studies, have each referred to a "disorganised and dilatory" factor, while Taylor (1984), from a different perspective, refers to what may be construed as the present three, and four others, including a "social" approach. The three referred to above are the more important ones, and are the focus of the present research.

(2) The names "surface", "deep", and "achieving" are suggested as these now have wide currency in the literature on student learning (in the original model, the first two approaches were referred to as "utilising" and "internalising" respectively). It should be noted that deep and surface approaches are different in kind from the achieving approach. The strategies involved in the first two describe ways in which students engage the actual content of the task, while the achieving strategy describes the ways in which students organise the temporal and spatial contexts in which the task is carried out. There is, then, no inconsistency in rote learning in a highly organised way ("surface-achieving") or reading for meaning in an organised way ("deep-achieving"). It would, however, be difficult to see how one could simultaneously rote learn and seek meaning (which is not to say that these strategies may not be deployed successively, as they are for instance by actors when learning, then interpreting, their lines).

(3) Approaches to learning have been used to refer both to personality characteristics and to situation-induced reactions. Individuals are predisposed by their personality to adopt one approach in preference to another, while certain situations encourage or inhibit particular approaches: this trait-state interaction is explicated further in this paper.

(4) The three approaches are likely to lead to different qualities in the learning outcome; as explicated further below, surface leads to retention of factual detail and deep to structurally complex and affectively satisfying outcomes.

(5) Optimal results are likely when the strategy used is congruent with the student's prevailing motivational state: that, for example, a deep strategy works most effectively when the student is intrinsically motivated, or an achieving strategy when achievement motivated.

THE LEARNING PROCESS COMPLEX: THE LPQ AND SPQ

The learning process complex mediates between presage and product variables, and is operationalised by two instruments: the Learning Process Questionnaire (for secondary students) and the Study Process Questionnaire (for tertiary students). The bulk of the work to be reported below was carried out on the samples that were used to obtain norms for the two instruments. Full details of the sampling and procedures are in Biggs (in press). A brief summary appears below.

LPQ

The sampling was conducted in conjunction with another project of the Australian Council for Educational Research, and involved two-stage cluster sampling of all Australian children aged 14, and all Australian children in Year 11 (Lower Sixth) on specified dates in 1979, resulting in both samples with conservative error bounds of ±5 per cent for LPQ item types. For present purposes, it is convenient to designate these as the 14+ and 16+ samples, respectively. The final Ns for the 14+ sample were 644 boys and 709 girls, and for 16+, 462 boys and 517 girls.

Other data collected along with the LPQ responses (as relevant to the present study) included:

(*a*) Personal information such as sex, date of birth, educational and occupational aspirations, parents' occupation and education, favourite subject in school, ethnicity.

(*b*) Word Knowledge (WK) (Thorndike, 1973), which was a 40-item test of verbal ability, modified for Australian conditions (Keeves and Bourke, 1976).

(*c*) Locus of Control (LC), which was a 7-item short form of Rotter's scale, prepared for survey use by Lokan *et al.* (1982), with a high score indicating internal LC. The alpha coefficient of the shortened version is 0·75.

(*d*) Academic performance, which consisted of two self-ratings, each on a 5-point scale, of performance as compared with peers (SRP) and satisfaction with current performance (SAT). In the case of a subgroup of 320 students from one State, external examination results were made available at the end of the following year.

SPQ

The sampling of tertiary students, from each of the advanced education (CAE) sector and the university sector, could not be randomly carried out as in the case of the LPQ. Rather, the writer had to rely on the goodwill of colleagues and contacts to distribute the questionnaires in class, and to collect and send on completed returns. The attempt was made to sample from faculties represented in both sectors (Arts, Education and Science). There was a total N of 823 university students and 1550 CAE students, representing 5 and 10 institutions respectively.

The following information was obtained with the SPQ responses: Faculty or school, full or part-time, sex, date of birth, ethnicity, educational plans; and as for the LPQ, the self-ratings SRP and SAT. For policy reasons, the forms had to be administered anonymously and so "hard" data on performance were not available.

LPQ/SPQ scales and subscales

The 36-item LPQ and the 42-item SPQ are each scored in the same way, by summing responses to the 5-point Likert items comprising each subscale. There are six scores, three motives and three strategies. The sum of the related motive and strategy subscales yields the approach scale score, as in Table 2.

TABLE 2

Structure of the LPQ/SPQ Scale and Subscale Scores

Level	Surface		Deep		Achieving	
Subscale	Motive	Strategy	Motive	Strategy	Motive	Strategy
Scale	Approach		Approach		Approach	
Composite			Approach			

The reliabilities, suggested uses, and norms of the subscales, scales, and the composite score are fully discussed and presented in Biggs (in press). The Australian norms are expressed in deciles and are available separately for sex and year (LPQ) and sex, institution, and faculty (SPQ).

These and other associations will be presented here as they are of substantive interest to this paper.

With regard to sex, males scored higher than females on Surface Approach in both secondary and tertiary samples; and females higher than males on the Achieving Approach in the tertiary, but not secondary, samples.

Age differences are discussed below (see Table 3 and discussion).

Institutional and faculty differences in the tertiary sector indicated that CAEs were highest on Surface, and universities on Deep, Approaches. Science students scored higher than Arts on Surface, and Arts higher than Science on Deep, independently of institution type. These main effects were slightly modified by interaction between institution and faculty as discussed in Biggs (1982) and Biggs and Kirby (1983).

There were several other variables that were directly associated with Approach scores. These are summarised in Table 3, as they give some idea of what these scores might be measuring.

TABLE 3

SUMMARY OF EFFECTS ON LPQ AND SPQ APPROACH SCORES

	Performance					Experiential			Ind. Diffs.	
Approach	SRP	SAT	Exam	Structure	Educ. Plans	Age	ESL	Home (4)	Ability	ILC
LPQ: Surface	−	0/−	−/− −	−	−	−	0	− −	−	−
Deep	+	+	(1)	+	+	(2)	+	+	0	+
Achieving	+ +	+ +	+	0/+	+	(3)	+	+	0/−	−
SPQ. Surface	−	0	(5)	−	−	−	0	−		
Deep	+	0	(6)	+	+ +	+ + (7)	+	(8)		
Achieving	+	+	(6)	0	+ +	+ (7)	0	−		

Notes: (1) + in "Favourite subject" only.
(2) − for boys, + for girls.
(3) − for boys, 0 for girls
(4) Extent of Father's Education.
(5) − − in Science, 0 in Arts.
(6) +/+ + in Arts, 0 in Science.
(7) Non-linear: decline from ages 18 to 22, sharp rise thereafter to age 40 years and over.
(8) Curvilinear. deep highest with least education and with most.

+ and − indicate direction of effect: +/− = significant but relatively weak; + +/− − = relatively strong

Performance

Correlations with self-rated performance relative to peers (SRP) are consistent across both samples: Surface correlates on average around −0·15 (which is highly significant statistically, given the Ns involved), with Deep correlating positively in the low 0·20s and Achieving in the 0·30s. Approach was less consistently related to satisfaction (SAT) although the direction was the same.

Hard performance data were available only for a section of the secondary sample, and that was fifteen months after the LPQ had been administered. Surface correlated negatively, and Achieving positively, with public examination results (both around 0·20 to 0·30) while Deep correlated positively only in the student's favourite subject. This last finding is in keeping with the congruence hypothesis, as the deep approach would be expected to be deployed only in the subjects in which the students were intrinsically motivated. Corresponding data for the tertiary samples were not available, but Watkins and Hattie (1981) independently report correlations between SPQ Surface Approach scores with first year Science of −0·40, and Deep and Achieving with first year Arts and Economics around 0·30.

The quality of performance, as reflected in the structural complexity of students' open-end responses to stimulus questions, may be assessed by the SOLO Taxonomy (Biggs and Collis, 1982). Studies relating this aspect of performance to approaches to learning have been carried out amongst secondary (Kirby and Biggs, 1981) and tertiary (Biggs, 1979) students. Deep Approach was clearly implicated in complex responding by both groups; Surface with low complexity of response but a corresponding high recall of factual detail; and Achieving affected performance according to the student's perception of what constituted success. In the 1979 study, students seemed to trade structural complexity of response for fidelity of recall of detail, according both to instructions and to predisposition.

Educational plans

At both secondary and tertiary levels, Deep and Achieving Approach Scores are closely associated with plans for furthering one's education as far as possible, and Surface with ending formal education early: at age 14 by leaving school as soon as legally possible, and at university and college as soon as the present qualification is completed.

Experiential

Three variables relate to general background experience. Age itself correlates negatively with Surface in both secondary and tertiary groups. In school, girls' Deep scores increase with age, but boys' decrease, as also do boys' Achieving scores. These age relationships may be associated with perceptions of schooling; if so the implications for boys are disturbing.

At the tertiary level, both Deep and Achieving decrease from first to third year. This is also a disconcerting effect and it has been independently replicated (Watkins and Hattie, 1983). It probably reflects both "institutionalising" effects on pass-level students, and insecurity about graduate employment. In fact, both Deep and Achieving Approaches reach a minimum at age 22. Thereafter, both increase monotonically with age, until age 40 at least, Deep more so than Achieving.

Another non-institutional experiential factor which consistently related to the Deep Approach was ethnicity, specifically in the extent to which English was a second language to the student. It seems that the experience of bilinguality, with its ongoing search for clarifying meaning and monitoring one's own verbal output, facilitates the development of a deep approach. It would seem very plausible that this relationship would be mediated by a metacognitive process.

Finally, extent of father's education is associated with low Surface and high Deep and Achieving scores, with the slight modification that a Deep Approach in tertiary students is associated with both very low (primary) and high levels of father's education.

All these results are discussed in more detail in Biggs (in press).

Individual differences

Individual differences have been investigated in secondary students only, and, as might be expected on the basis of the original model (Figure 1), some quite complicated interactions with approaches to learning have been found, the more important of which will be discussed below in a later section. The first order correlations with learning approach and verbal ability (WK) and locus of control (LC) are nonetheless interesting.

Ability correlates negatively with Surface scores, as might be expected on the basis of performance relationships, but zero with Deep and Achieving scores: deep students are not necessarily more able, despite their better performance. Indeed, at 16+, there is a small but significant *negative* correlation between WK and use of the Achieving Strategy: it is possible that at this level of schooling, self-organisation and time management are what the less able, rather than the more able, tend to do in order to survive. Internal LC, on the other hand, consistently correlates negatively with Surface and positively with Deep scores.

Summary

To generalise from these findings, a surface approach is associated with poor academic performance, as assessed objectively and in terms of students' self-ratings, with poor motivation, but with good recall of factual detail when that is appropriate. In line with this, students endorsing a surface orientation in general intend to get out of formal education early, have an external locus of control, low verbal ability, and parents (particularly the father) of low education.

A deep approach is associated with good academic performance, whether assessed in subjective, objective, or qualitative terms. Students using this approach have a good academic self-concept: they rate themselves favourably in terms of performance level and satisfaction with their level of performance, and intend to complete Year 12 (18-year-olds) in the case of secondary, and post-graduate work in the case of tertiary, students. A deep approach also appears to be facilitated by certain kinds of experience: favourable home background, the experience of bilinguality, and the planning and decision-making that is part and parcel of adult living. It seems significant that the individual difference variable associated with the development of a deep approach is not intelligence *per se*, but an internal locus of control.

An achieving approach is associated with many of those associated with deep, but less strongly with the experiential factors (age in adulthood, English as a second language, and home background) and not at all with structural complexity of performance.

It seems that one may set up a hierarchy of approaches to learning on the basis of their relationship (at least) with complexity of outcome. Deep would be highest, leading to the most complex outcomes, surface the least personally involving and leading to low structure-high factual outcomes, and achieving in between these two but nearer to deep.

The analysis so far paints only a very broad picture, from which several features are missing:

(*a*) How do personal, situational, and process variables interact to affect outcome?

(*b*) How can one explain the dynamics of this interaction?

ASPECTS OF METALEARNING

The dynamic link between personal, situational, process, and outcome variables is suggested to be a metacognitive process, based around the notion introduced above of motive-strategy congruence. This may be understood in terms of the *personal study contract* (Taylor, 1984) that a student makes with himself or herself. The contract in effect acknowledges the intentions and purposes held on arriving at university, and the actions one is thereby committed to if those intentions are to be realised (with allowance for some revision as experience modifies both what is desirable and what is practicable). As Taylor says: "To make a study contract that has a reasonable chance of succeeding . . . students need to be aware of their own abilities in relation to the situational context" (p. 254).

Whether one is talking about motive-strategy congruence, or a personal study contract, the above discussion fits squarely with current definitions of *metacognitive* processes as "one's knowledge concerning one's own cognitive processes and products . . . (and) the active monitoring and consequential regulation of those processes in relation to the cognitive objects or data on which they bear" (Flavell, 1976, p. 232. See also Brown *et al.*, 1983; Brown, 1984). Although there are several, and confusing, meanings attached to the prefix "meta", essentially metacognition is a second order construct. Whereas the object of cognitive activity is a problem, datum, or suchlike, that of metacognitive activity is the cognitive process itself, not the object of that process.

The notion of motive-strategy congruence implies, first, that students are indeed able to enter that "metamotivational state" (Apter, 1982) of being able to interpret their own motives, and then to be realistically aware of their own cognitive resources in relation to task demands. They then need to plan, monitor, and control those resources: that is, to be metacognitive. It thus seems likely that the literature on metacognition would be of direct relevance to the study of students' study processes.

In view, however, of the wide range of mental events that is, or could be, covered by the term metacognition (Brown *et al.*, 1983), the term *metalearning* is proposed here for the rather specialised application of metacognition to the area of student learning. Metalearning, then, is a subprocess of metacognition (in much the same way as are metamemory and metamotivation) that refers specifically to learning and study processes in institutional settings, and more particularly to students' *awareness* of their motives, and *control* over their strategy selection and deployment.

In the present paper it is intended to present findings of recent studies that bear upon these two phases in metalearning, that is, of being aware of the available options, and of exerting control over those options. Next, interactions with performance are described, with particular attention to individual differences and motivational state. A model of metalearning emerges, which has implications both for intervention strategies for students in difficulty, for teaching, and for further research.

On being aware of the options

In the present model, six parameters in the learning process complex are suggested: three motives and three strategies, which logically reduce to three dimensions, with a motive and a related strategy comprising each dimension. In other words, a factor analysis of the six subscale scores should reduce to three orthogonal dimensions if students were accurate in discriminating possible motives for learning and associating them with their appropriate strategies.

Biggs and Kirby (1984) subdivided a group of 320 Year 9 students (14-year-olds) into four different patterns of ability: the abilities used were simultaneous and successive processing (Das *et al.*, 1979), which for present purposes may be described as "reasoning" and "memory" respectively. Four subgroups were formed by splitting these two variables at their medians: students high on both reasoning and memory, those high on reasoning but low on memory, those low on reasoning and high on memory, and those low on both. Such a grouping is fairly drastic, and one would expect that, if LPQ subscale scores (the six motive and strategy scores) were factor analysed within each group, the factor structure would differ between groups. Using principal components with varimax rotation, and taking eigen values of unity or more, this is in fact what happened (Table 4).

As can be seen, the *low-lows* do not make any discriminations at all, either between motive and motive, or even between strategy and motive; the learning process complex must seem opaque and meaningless to them. The *low*-reasoning/*high*-memory group made clear discriminations, putting motive with strategy to form a two-dimensional learning process complex: they ally the Surface Approach with the Achieving Approach to form a composite Surface-Achieving, as might be expected of a group biased towards memory over reasoning, while Deep and Achieving Strategy form a factor of their own that is unrelated to wanting to achieve. The *high*-reasoning/*low*-memory group do the opposite. They, too, perceive a two-dimensional space, but associate getting-on (achieving) with the Deep Approach, as befits their bias towards reasoning.

The students *high* on both abilities perceive all three dimensions, basically in line with the "logical" model.

One might make two points about these discriminations within the learning process complex: the brighter groups make *more* discriminations than the duller, and the line-up of strategy with motive accords with the cognitive "bias" of the students concerned.

TABLE 4

FACTOR STRUCTURES OF LEARNING PROCESS COMPLEX IN 14-YEAR-OLD STUDENTS OF DIFFERING ABILITY PATTERNS

			REASONING						
			Low			High			
			1. N=90			3. N=72			
			I			I	II		
	Surface	M	65			—	78		
		S	49			—	85		
Low	Deep	M	·80			74	—		
		S	72			71	—		
	Achiev	M	66			65	—		
		S	80			69	—		
	Variance	%	48·3			35·2	23 9	59·1	
MEMORY									
			2. N=70			4. N=89			
			I	II		I	II	III	
	Surface	M	78	—		—	52	62	
		S	69	—		—	—	86	
High	Deep	M	—	74		84	—	—	
		S	—	85		81	—	—	
	Achiev	M	60	—		—	87	—	
		S	58	62		52	58	—	
	Variance	%	30 1	30·0	60·1	27·3	24·0	19·5	70 8

Decimals, loadings <0·40 omitted

In short, the first component of metalearning, being aware of one's motives and how one might use different strategies to realise them, shows a progressive increase in the four groups. Metalearning would appear to be least viable for the low-lows, most for the high-highs, while that for the others will be biased by their particular pattern of abilities.

Internal locus of control seemed to be implicated in the development of deep and achieving approaches to learning and this is easy to understand. The perception of meaning, having sufficient discipline over oneself to manage time appropriately, and being aware of one's motivational state, are all activities that implicate an inward-looking or introspective stance. There is also a considerable literature on this variable that links it effectively with learning in school (e.g., Gammage, 1982) and university (Biggs and Das, 1973).

It is worth asking, then, whether the sort of analysis of options described above would replicate using LC rather than ability patterns as the independent variable. In order to obtain extreme groups of adequate size, the 14+ and 16+ samples were divided into top (internal) and bottom (external) 20 per cent of the LC distributions, and the six LPQ subscale scores were likewise subjected to principal components analysis with varimax rotation. The 16+ groups produced very similar two-factor solutions (Surface and Deep-Achieving) as did the internal 14+ students. The 14+

externals, however, produced a single factor exactly like the low-low abilities group (Table 5).

Similar analyses of the extreme 20 per cent of the verbal ability (WK) distributions at 14+ and at 16+ were conducted but the results were the same across all groups: the two-factor Surface and Deep-Achieving solutions.

TABLE 5

PRINCIPAL COMPONENTS ANALYSIS, WITH VARIMAX ROTATION, OF LEARNING PROCESS COMPLEX IN HIGH INTERNAL AND HIGH EXTERNAL LOCUS OF CONTROL STUDENTS AT 14+

	Internal LC N=286			External LC N=271
	I	II		I
SM		76		65
SS		81		48
DM	78			74
DS	83			80
AM	71			82
AS	77			74
Total %	42 1	22·4	64 5	50 9

Decimals, loadings <0·40 omitted

It thus seems that the first stage of metalearning — recognising what strategy goes with what motive — is generally present in most students by 16 years of age, and in internally controlled 14-year-olds, but is not present at all in 14-year-olds of low ability in both memory and reasoning (Biggs and Kirby, 1984), or who are of average ability but are not given to the kind of introspection that leads them to reflect on their own cognitive processes.

On the development of control

The second stage of metalearning is the exercise of control over one's perceived strategic options. Brown *et al.* (1983) have distinguished several specific components of the regulative aspect of metacognition: planning, which itself subdivides into metaplanning, executive decision-making, world knowledge decisions, plan-abstraction and specific plan decisions; monitoring; and checking. Certain of these aspects, particularly of planning, would appear to be of direct relevance to the question of executive control in metalearning, but a superordinate factor over-riding specific mechanisms as such would again implicate the construct of internal locus of control. The *belief* in one's ability to exercise control over one's own learning is prerequisite to the deployment of particular aspects of planning or checking.

Given, then, the differences in the LPQ factor structures as mediated by LC and age, as discussed above, one might ask what age-related differences might be found in the construct of LC itself. One way of looking at this is to factor analyse the items themselves that make up the present LC scale, within each age level (Table 6).

At 14+, the construct of LC seems to be poorly articulated: only one factor emerges, with only 41 per cent of the item variance accounted for. One item (No. 4) is worded in the opposite direction to the rest and simply disappears, rather than loading negatively as it does by Year 11. The construct of LC seems unstable in 14-year-olds if a complication in format of one item destroys its meaning.

At 16+, two factors emerge, accounting for 52 per cent of the variance: (lack of) confidence in one's control over significant life events, and belief that one's life is

TABLE 6

PRINCIPAL COMPONENTS ANALYSIS, WITH VARIMAX ROTATION, OF THE SEVEN LOCUS OF CONTROL ITEMS

	14+ (N=1350)	16+ (N=980)		
	I	I	II	
1. Good luck is more important than hard work for success	74		84	
2. There is no sense in making plans, they usually don't work out	68	62		
3. Every time I try to get ahead, something or someone stops me	62	69		
4 It's worth having a goal in life to work for		−52		
5. Getting ahead in a job depends more on good luck than hard work	70		84	
6. People like me don't have much of a chance to be successful in life	67	71		
7. When people are born, the success they are going to have is already on the cards, so they may as well accept it and not fight against it	71	51	37	
	41%	28%	24%	52%

Decimals and loadings <0·35 are omitted

controlled by the external factor of luck. (The method of scoring reverses the wording so that a high score indicates *internal* LC). Item 7 refers both to control and to luck, and loads on both factors.

When the factor scores are correlated with other variables, their meaning becomes clearer (Table 7).

The 14+ factor has strongest correlations with ability, then with performance, and with Surface, Achieving and Deep Approaches in that order. The self-control factor at 16+ has relatively weaker correlations with ability than at 14+, similar ones with performance, but stronger ones with Deep and Achieving. The rejection-of-luck factor has a marginally stronger correlation with ability, but weak to zero ones with other variables. While the relationships between LC, approach to learning, and performance may be attributable to ability at 14+, they may be explained instead at 16+ by the development of the metacognitive component of self-control.

If these findings are put together with those reported in the previous section, a composite model implicating ability and internal LC in the development of metacogni-

TABLE 7

CORRELATES OF LOCUS OF CONTROL FACTORS*

			Performance		Approaches		
Sample	LC Factor	Verbal. Ability	SRP	SAT	Surface	Deep	Achieving
14+ (N=1350)	Global	0·34	0·26	0·17	−0·21	0·08	0·14
16+ (N=980)	Personal Control	0·16	0·23	0·14	−0·21	0·17	0·23
	Reject Luck	0·19	—	—	−0·10	—	0·08

* Shown only if P<0·05

tion emerges. As 14-year-olds begin to discriminate and to match strategies with motives, metacognitive control also emerges. Two potentially important conclusions may be drawn. First, 14-year-olds typically do not appear to distinguish luck from a belief in their own executive control whereas 16-year-olds do. Second, the relationships between LC, approach to learning, and performance are mediated more by *cognition* (ability) in 14-year-olds, and by *metacognition* (LC controlling approach to learning) in 17-year-olds. Some of these interactions are explicated in the next section.

Interactions with performance

1. *Ability and locus of control*

The work summarised to date suggests that approach to learning is likely to interact with individual difference variables in so far as the latter index the student's capability for metalearning. One would expect, for example, that a student with internal LC would use a deep approach more effectively than a student with external LC, while the latter might be able to study more efficiently by using the less metacognitive achieving approach to structure time schedules more effectively.

Such expectations may be examined for those students whose HSC (an external examination at 17+) results subsequently became available. Several ANOVAs were carried out involving ability (WK), LC, and in turn the three learning approaches split at the median as independent variables, and with Aggregate (sum of best four HSC examination marks) and some individual subjects as dependent variables. Several significant interactions were found, two of the most interesting being reported in Table 8 and Figures 2 and 3.

FIGURE 2

EFFECTS OF ABILITY, LOCUS OF CONTROL, AND DEEP APPROACH ON HSC AGGREGATE

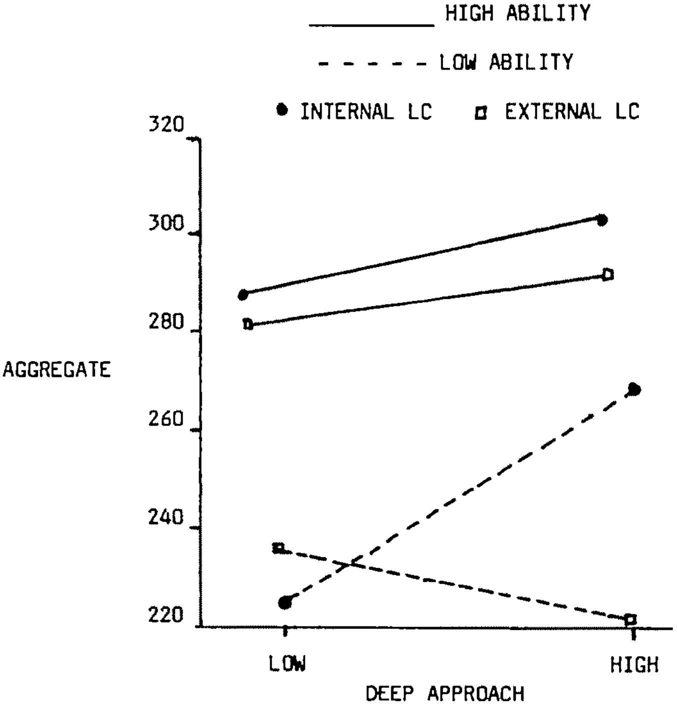


TABLE 8

ANOVAS OF ABILITY × LC × APPROACH (DEEP OR ACHIEVING) ON HSC AGGREGATE

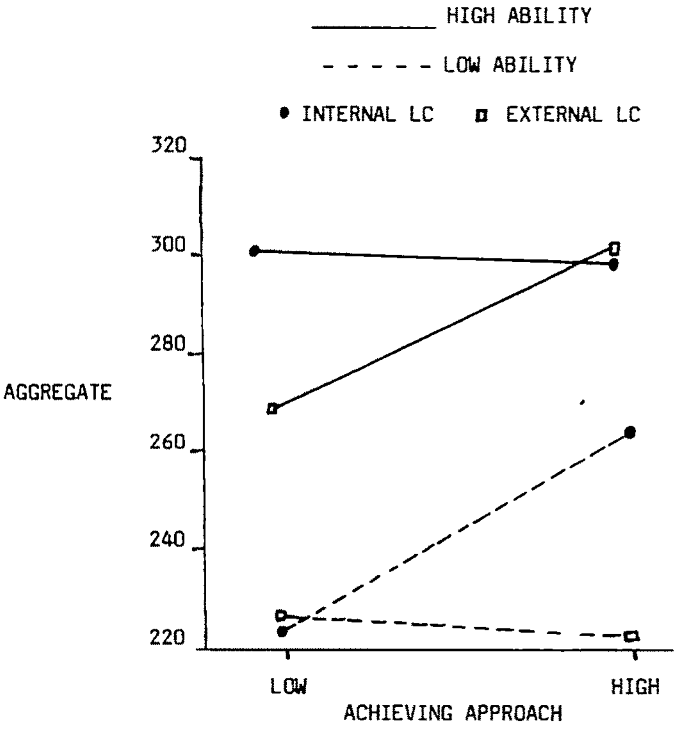
P<	Deep	Achieving
Verbal Ability (WK)	0·000	0·000
Locus of Control (LC)	(0·10)	(0·10)
Approach (Deep or Ach)	0·01	0·001
WK × LC	—	—
WK × A	—	—
LC × A	0·05	—
WK × LC × A	(0·10)	0·05

Basically, the LC × Approach interaction (Figure 2) suggests that the Deep Approach works more effectively with internally controlled students independently of ability but as the second order interaction is marginally significant, and, as a glance at Figure 2 shows, the Deep Approach does not work at all with low ability externals, who would be expected to be low on metalearning capability. It is noteworthy that low ability internals using the Deep Approach appear to gain over 50 Aggregate marks, performing nearly as well as high ability externals. Clearly, locus of control does indeed appear to be mediating the operation of the Deep Approach.

Corresponding data for Achieving Approach are illustrated in Figure 3.

FIGURE 3

EFFECTS OF ABILITY, LOCUS OF CONTROL, AND ACHIEVING APPROACH ON HSC AGGREGATE



In the case of low ability students, LC again appears to be mediating the approach to learning, with internals gaining 40 Aggregate marks. The picture with high ability students is, however, different from that illustrated in Figure 2. High ability internals appear to be working at a high level consistently, and independently of the Achieving Approach. This time the Achieving Approach appears to be facilitating the performance only of the externals, as if the internals are already controlling their own learning and are able to perform at a level that the externals can only reach with the specific prop of self-consciously managing their time and organising their work.

Data such as these explain why first order correlations between approach and performance are often relatively small, and sometimes non-significant: the same approach often does different things for different students. More importantly, the ability × LC interactions with approach implicate metalearning as at least one construct that helps explain these interactions, with perhaps different aspects of metalearning ability being involved with the deep and achieving approaches.

2. *Motive-strategy congruence*

The question of motive-strategy congruence raises two issues: the extent to which students who endorse a particular motive tend also to endorse the cognate strategy; and the extent to which congruent motive-strategy combinations are more effective than non-congruent ones. The first, weaker, aspect of congruence is easily settled. The correlation between any motive and its cognate strategy is consistently higher than that between that motive and any other strategy (Watkins, 1982; Biggs, 1984; O'Neil and Child, 1984).

The strong aspect of the congruence hypothesis is that such congruent combinations (which of course comprise the approaches) are more *effective* than non-congruent ones. The problem then becomes in part one of defining "effectiveness". If effective means in relation to the student's own personal goals, one simply cannot generalise. Indeed, the idiographic argument if pushed too hard becomes circular: congruent motives and strategies are those perceived by the student to be effective for his or her own purposes. If effective means in relation to independent criteria such as examination results, congruence results in effective performance only under certain conditions, to be described below.

While Watkins (1982) found no relationship between discrepancy scores between motive and related strategy subscales and performance, that study does not really address the crucial issue: how does a strategy relate to performance independently of the motive in which it is usually embedded, or in conjunction with a different motive? To answer that question, the three motive subscale distributions were split at their medians, for each of the 14+, 16+, CAE, and University samples, yielding eight motivational subgroups, ranging from *low-low-low* on all three motives through all combinations to *high-high-high* (see Biggs, 1984, for further details). Each strategy was then correlated, within each motivational subgroup for all four samples, with self-rated performance (SRP) and satisfaction (SAT). The results are quite complex, as there were 192 correlations between strategies and performance. Three-way ANOVA of the data, with Surface, Deep and Achieving Motives independent, and SRP and SAT dependent, showed strong main effects and interactions indicating that certain motive groups were more effective than others. It was also found that the rank order of effectiveness did not differ between SRP and SAT.

The eight motivational subgroups, in order of effectiveness, and with a brief description of each, are given in Table 9, together with the number of statistically significant correlations between strategy and performance that involve congruent or non-congruent motive-strategy combinations.

TABLE 9

NUMBERS OF SIGNIFICANT CORRELATIONS BETWEEN STRATEGIES AND PERFORMANCE FOR EIGHT MOTIVATIONAL SUBGROUPS

	Motivational Subgroup (Rank order Effectiveness)				No. Sig. Effects		
	SM	DM	AM	Description of Subgroup	Cong	Nil	Non-cong.
1	L	H	H	Deep Achievers	7	17	0
2.	L	L	H	Competitive achievers	4	20	0
3.	H	H	H	Defensive achievers	2	16	6
4.	H	L	H	Surface achievers	3	19	2
5.	L	H	L	Deep idealists	5	14	5
6	L	L	L	Unmotivated	4	11	9
7.	H	H	L	Defensive actualisers	0	14	10
8.	H	L	L	Under-achievers	0	19	5
Total No. Strategy Effects					25	130	37

The majority of strategy/performance correlations (130 out of 192) was non-significant, and out of the 62 significant ones, 37 involved strategies associated with non-congruent motives, 25 with congruent. The evidence for the congruence hypothesis is thus not very strong. Table 9 does show an interesting interaction, however, in that the two most effective motivational groups ("Deep achievers" and "Competitive achievers") contain 11 congruent effects and no non-congruent ones, while the two least effective groups ("Defensive actualisers" and "Underachievers") contain no congruent effects and 15 non-congruent ones. That difference is well beyond chance.

In other words, the well-motivated and high achieving students tended to select strategies that were congruent with their motivational state and to deploy them effectively. Poorly motivated students tended to select strategies that were not congruent with their motives, yet sometimes those strategies (particularly Achieving) were effectively used. Congruence, or rather non-congruence, did not seem to be an issue with these students, in either selection or deployment. In such cases, "strategies" appear to be used with little or no metacognitive involvement on the part of the students; they appear rather to be functioning in the same way as "techniques" (Brown, 1978) or "tactics" (Snowman, in press), i.e., as short-term props to learning that do not involve any metacognitive insight on the part of the learner.

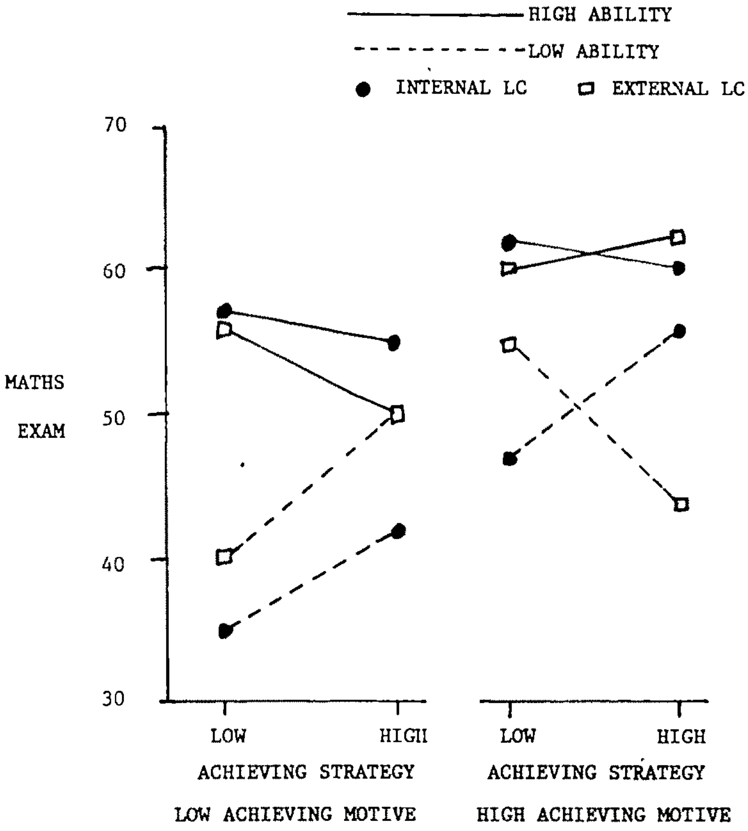
Two quite different mechanisms seem to be involved in explaining the correlations at each motivational extreme. Strategy deployment by well motivated and achieving students seems to be based on a high level of metalearning and usually is effective. At the other motivational extreme, strategy use parallels that of a technique or tactic, and is done by following procedures involving little or no metacognitive awareness.

Further illustration of these two types of strategy use is provided by other ANOVAs of the 16+ HSC results. It will be recalled (Figure 3) that the Achieving Approach interacted with WK and LC to affect the Aggregate differentially. In the following analysis, Achieving Approach was divided into its Motive and Strategy components, so that a four-way LC × WK × AM × AS analysis could be carried out on the results. Results were clearest with Mathematics marks dependent (Figure 4).

The four-way interaction depicted above ($P < 0.05$) indicates conditions under which congruence is present or absent, and when strategy use is effective or not. Let us first take High Achieving Motive students (right-hand graph). The bright students all do well, regardless of LC and of Achieving Strategy, and better than any other group of students (there is no difference between the four subgroups here). There is, however, a

FIGURE 4

EFFECTS OF ABILITY, LOCUS OF CONTROL, ACHIEVING MOTIVE AND ACHIEVING STRATEGY ON MATHEMATICS PERFORMANCE



strong disordinal interaction amongst the less able: internals make very effective use of the Achieving Strategy but externals do correspondingly better without it.

The Low Achieving Motive students show quite a different pattern. Regardless of LC, able students do worse using the Achieving Strategy, which fits with the congruence hypothesis as they are inadequately motivated for it, but the low ability students, again irrespective of LC, do better, which is certainly not in conformity with the congruence hypothesis.

These data, and particularly those describing when the Achieving Strategy is helpful or harmful, are explicable in terms of the extent of metalearning involved. First, the Achieving Strategy is not an issue with high ability, highly motivated students; their high performance is likely to be maintained by other strategies, particularly Deep. It is, however, an issue with lower ability students, and with high ability students who are poorly motivated. Low ability, highly motivated internals use the strategy very effectively; they appear to be making appropriate use of metalearning. Low ability highly motivated externals, and high ability poorly motivated students, whatever their locus of control, cannot: the Achieving Strategy in these three groups seems to get in the way, as if they are metacognitively developed enough for the strategy to be an issue, but not developed enough to use it appropriately.

Finally, low ability, poorly motivated students, who might be reckoned to be the lowest in metalearning ability, nevertheless seem to use the Achieving Strategy effectively. It seems that here, as in the lowest motivational groups in Table 9, it does seem to help as a technique or tactic.

Of course, these complicated patterns are being interpreted *post hoc*. One almost certainly could not have predicted them on the basis of metalearning theory. Having obtained them, however, one is able with the wisdom of hindsight, and with the support of a number of highly compatible observations, to paint a picture that makes considerable theoretical sense.

Summary

The data reviewed here suggest a developmental account of the role of metacognition in students' approaches to classroom learning and studying:

(1) Awareness of one's metamotivational state (Apter, 1982), appears to be the first stage, and it involves awareness of what one wants to get out of a learning situation or a particular task, followed by selecting strategies appropriate to those self-defined ends. This general stage seems to be available to many 14-year-olds, but not to those of low ability in general or of normal ability but a high external LC. It does not necessarily mean, however, that those strategies will be deployed effectively.

(2) Between the ages of 14 and 16, approximately, students seem to articulate more clearly what is meant by exerting control over their lives, which seems in the context of institutional learning to become related to a greater awareness of and increasing control over approaches to learning.

(3) Even at the upper end of the secondary school, however, many students do not appear to have the metalearning capability to use learning strategies appropriately. Following are some of the factors that appear to set the stage for academic self-awareness and successful strategy deployment:

(*a*) A relatively high and appropriate spread of both reasoning and memory.

(*b*) A rich content background that encourages planning and self-awareness, such as seems to be provided by bilinguality, and by living and planning adult experiences.

(*c*) Internal LC.

(*d*) A high degree of the appropriately orientated motivation, which in the academic context refers to intrinsic interest and to achievement motivation.

In short, effective self-direction in learning is a complex and sophisticated process that requires certain cognitive and affective personal characteristics, a suitable background of experiences, and a facilitative environment context. If we put these conclusions together with those deriving from the preceding section a rather more precise version of the model of student learning outlined in Figure 1 emerges.

AN ELABORATED MODEL OF STUDENT LEARNING

The work reviewed above provides a reasonable basis for filling in some of the detail in Figure 1. The three levels of presage, process, and product are elaborated, and then their inter-relations described.

Presage variables

Personal factors include levels and patterns of certain cognitive abilities, locus of control, and character of background experience, which includes home background, bilingualism, and mature age experience.

Situational factors that have emerged in the present research programme include the nature of the task, as for instance in differences relating to Arts and as opposed to Science; type of institution, as in university/CAE differences; and instructions to attend to meaning or to factual detail. Other investigators have looked at such situational pressures in students as assessment (Fransson, 1977) and student perceptions of institutional requirements (Laurillard, 1979; Ramsden, 1979).

Process variables

The learning process complex is envisaged as comprising several motives and strategies, three of each being operationalised by the LPQ/SPQ. Motives are considered as being prior to strategies, and as springing in part from the individual's personality structure and in part from situational pressures. Strategies are envisaged as arising out of motivational states in accordance with task demands. Three main approaches are considered: deep, achieving and surface, and each has its own motive-strategy combination.

Deep approach has the closest links with personality factors, is associated with the most complex and most satisfying learning outcomes, and requires least instructional structure or formal teaching support. Achieving approach is slightly "further" from the personological and closer to the situational, and works well with more instructional support structures particularly those emphasising competition. Nevertheless, achieving may overlap with deep to form an institutionally adaptive deep-achieving approach. Surface approach is the most susceptible to situational pressure, is associated with highly directive teaching, and leads to factually specific outcomes which are often associated too with negative affect. At the bottom of the process factors are the tactics or techniques that are highly task specific, with little or no transfer to other tasks, and that can be taught directly to students.

All three approaches may involve degrees of metalearning in that a self-conscious examination of the task, its context, and one's goals, may be involved in arriving at and deploying the most suitable approach to a given task. Metalearning is, however, most likely to be involved with the deep approach, so that in the search for meaning, the student is aware quite self-consciously of the "clicks" of comprehension and the "clunks" of incomprehension (Anderson, 1979). Such self-conscious awareness likewise characterises the achieving approach (see below, the Stanford study), but it is also true that the achieving approach may be used in an uncomprehending way (as is likely to have been the case of the low ability, poorly motivated students in Figure 4, and as revealed in many of the student comments collected by Tabberer, 1984). Self-conscious awareness may also occur in using the surface approach, as occurred for instance when students were instructed to learn fact and detail (Biggs, 1979), but it is more usually the case that the surface approach is deployed inappropriately, not as a result of a metacognitive decision but of habit or of despair.

Product variables

The various subjective, quantitative, and qualitative estimates of performance that have variously been used in the present research programme collectively suggest two major parameters of performance that covary:

(1) *Structure-Fact (S-F) Ratio.* All academic performances may be described in terms of (*a*) the extent to which the learning and correct reproduction of detail is paramount; and (*b*) the extent to which the structure in which the detail is embedded is paramount. There are in fact certain tasks in which reproduction of detail is considered to be important *per se*, rather than as being merely illustrative of the deeper structural features of the task. Such would be the learning of script and orthographics and lexicon particularly in the early stages of foreign language learning, or learning scientific/

mathematical formulae that are only partially understood if at all. In other tasks, structural complexity is more important, and a convenient way of operationalising this structure is suggested by the SOLO Taxonomy (Biggs and Collis, 1982) in which the relationships that inhere between the components of a learning outcome are taken to represent the quality of that outcome.

The inverse relationship between (*a*) and (*b*) is conveniently captured by the term Structure-Fact, or S-F, ratio. A task with a high S-F ratio indicates that a complex performance is required, where relationships between component data are to be emphasised; a task with a low S-F ratio would emphasise the fidelity with which specific data are to be recorded and reproduced, with little reference to the whole of which they are a part.

(2) *Affective involvement*. The affective outcomes of learning range from highly positive, as in intrinsic motivation or in expressed student satisfaction, to quite aversive. The latter reaction may occur either in the case of a complexly structured task, but which the student is unable to handle appropriately, so that the outcome is poorly structured; or in the case of a task with an inherently low S-F ratio.

Although these outcome factors are conceptually distinct, they covary in practice. To the extent that the student produces a high S-F ratio outcome, the affective involvement is likely to be high and positive. As the S-F ratio of the student's product (not necessarily of the task set) declines, affect likewise becomes increasingly disengaged, and a state of aversion may be reached. Clearly, there are task differences in S-F ratio, and individual differences in the extent to which a particular S-F ratio can be processed and in which involvement is aroused, but the concept of a two-factor performance outcome, comprising a structure-fact ratio and affect, is useful for describing process-outcome relationships in general terms.

The overall relationships between the three presage, process and product stages are expressed in Figure 5.

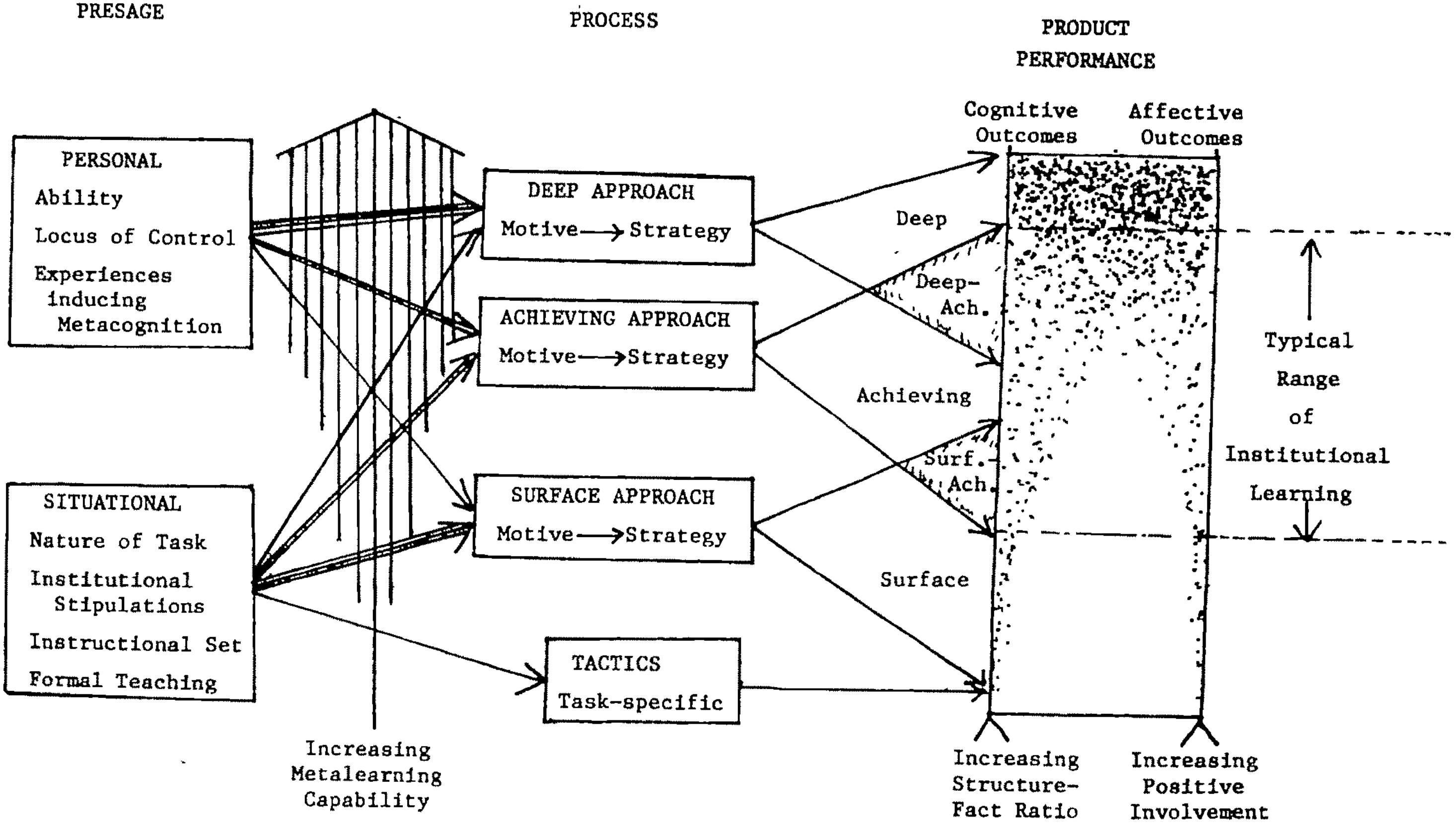
Figure 5 is meant to suggest the points that a deep approach is more closely tied to personological factors, and surface to situational, with achieving in between but closer to deep in most individuals. The arrows to performance indicate the likely range of S-F ratio and involvement targetted by the approaches. Overlap between deep and achieving represents the deep-achieving composite that fits well with the high structure end of academic performance, while the less common surface-achieving counterpart fits lower down the range of academic performance.

Metalearning is represented as increasing vertically. At the level of task-specific tactics metalearning is not involved at all: students do not need (or want) at this level to be aware of their own perceptions of means-end relationships, or even to know why they should be doing this and not that. The surface approach, too, is often an unaware gut-reaction to the task, as may sometimes be the case with achieving. It is, however, difficult to imagine the deep approach operating effectively by the upper secondary level in the absence of metacognitive awareness.

The performance domain depicts a vertically increasing ratio of structural-to-factual outcomes, and increasing positive affective involvement. A range of typical academic performance is suggested (relative to the level of institutional learning), which is largely encompassed by the overlap between deep and achieving approaches. A level of performance which exceeds that range, where the task is complex in the extreme and involvement, perhaps ecstatically high, might include fiction writing, say, or research activity. At the other extreme, there is a corresponding range of unstructured and reproductive tasks that may be legitimate in the psychological laboratory, but should be rare in the classroom. It may well be, however, and many

FIGURE 5

Elaborated Model of Student Learning



studies of student perceptions would indicate so (e.g. Marton *et al.*, 1984), that students wilfully or out of desperation define classroom tasks in this way.

IMPLICATIONS AND FURTHER RESEARCH

The elaborated model seems to have several educational implications. An immediate one, which is also in a sense a test of the model, is to do with strategies of intervention in order to improve student learning. We shall deal with this important point at some length, and then consider more general implications for teaching and research.

The problem of intervention

A long-standing issue in the student learning literature is the extent to which students can be trained to become better learners. Gibbs (1977) puts the question succinctly but with some scepticism: "Can students be taught how to study?" His answer is "No"; learning to him is a self-defined and self-discovered process that one needs to encounter and negotiate for oneself. Indeed, Tabberer (1984) cites many examples of secondary, even sixth form, students being untouched by, or actively resisting, the advice given in study skills courses.

The present model suggests that the answer would depend on the student's metalearning capability. If there are ways in which an outsider, such as a teacher or counsellor, can facilitate students' analyses of their own resources in relation to task demands, this would seem to be potentially very useful. After an extensive review of a similar problem, Wagner and Sternberg (1984) conclude that "Emphasis on metacognitive training does result in some degree of durability and transfer" (p. 199). Thus, if study skills courses — to call them that for the moment — can produce a level of self-awareness such that students can perceive what they want and how to get it, and they want it sufficiently, then it is likely that the students concerned could indeed become better learners.

If, however, that sort of self-awareness cannot be induced, for whatever reason, then it seems that one is left with teaching highly specific tactics that are close to the task, and "do the trick" without necessarily any metacognitive insight on the part of the student.

These considerations suggest a two-stage intervention model, in which the student's metalearning capability is taken as the point of departure for determining the style of intervention (Table 10).

There is a wealth of literature on the second, tactics, model of intervention, much

TABLE 10

AN INTERVENTION MODEL BASED ON STUDENTS' CAPABILITY FOR METALEARNING

Metalearning Capability	Approach to Learning	Type of Intervention
1. High to medium: students cognitively sophisticated and already well motivated or capable of becoming so	Deep-oriented. motive-strategy congruence important, and can engage tasks with a view to high transfer	Non-directive: students' metacognitive processes the target for intervention
2. Low: students with external LC, poorly motivated, and low abilities profile (probably)	Surface-oriented: motive-strategy congruence not an issue. Specific, low transfer, tactics linked to particular tasks	Directive: task-specific tactics and algorithms the target, not students' comprehension or self-monitoring

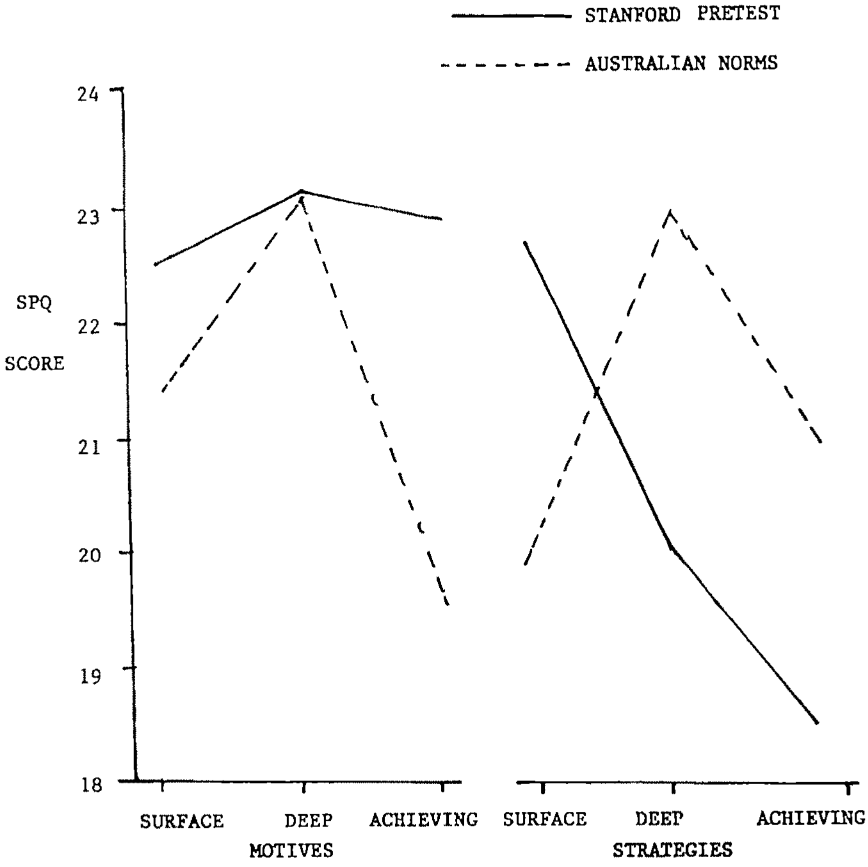
of it in the area of mental retardation (see Brown, 1978), but relatively little on the one presupposing a high level of metalearning. One example has, however, been in the literature for years: the work of Johnson Abercrombie (1969) with medical students. Her work presupposes a high degree of metalearning, and indeed her description of what she does is itself an excellent definition of metalearning:

> "My hypothesis is that we may learn to make better judgments if we can become aware of some of the factors that influence their formation . . . the student learns by comparing his observations with ten or so of his peers. He compares not only the results, but how the results were arrived at . . . What the student learns, it is hoped, is not only how to make a more correct response when he is confronted with a similar problem, but more generally to gain firmer control of his behaviour by understanding better his own ways of working" (pp. 18-19).

Another study, conceived within the framework of the present research, and which clearly brings out the issue of motive-strategy congruence, was conducted at the Learning Assistance Center, Stanford University. The Center provides a course for credit ("LAC-1 — Effective Learning Skills") for students who are dissatisfied with their grades. Fuller details of the course and context are provided in Biggs and Rihn

FIGURE 6

Stanford LAC-1 Pretest SPQ Scores Compared to Australian Norms



(1984): here, it just needs to be understood that Stanford has a highly competitive entry, expensive fees, and continual pressure to maintain high grades. Students who perceived themselves "at risk" in that context were already very bright and highly motivated, but were dissatisfied because they spent a lot of time in study and were not achieving high grades.

Two intakes in 1981 and 1982 were studied. A suitably reworded version of the SPQ was given to each group immediately before (pretest) and after (post-test) the nine week course, LAC-1. The pretest profile of motives and strategies is given in Figure 6, as compared to Australian norms. One would, of course, expect differences between such an exceptional group and the mean scores of unselected groups from another country; the present point is rather to compare the two sets of data with respect to motive-strategy differences.

A first glance shows that whereas the Australian motive and strategy profiles followed a similar outline, those of the Stanford group did not: Deep and Achieving Strategies fall well below their respective Motive profiles. More specifically, the Stanford students were the same as the Australian students on Deep Motive, and higher than the latter on Surface and especially Achieving Motives, but their strategies were surface dominated. LAC-1 students were very much higher than the Australian students on Surface, and correspondingly lower on Deep and Achieving. The picture is one of gross imbalance: the LAC-1 students seem highly fired-up motivationally, but with only a surface strategy to get them anywhere. Small wonder the students were so dissatisfied with their approaches to learning.

The course itself involved considerable self-direction in goal setting and monitoring of strategies, both through self-evaluation and through peer interaction. In conjunction with this, the usual time management skills (some of which are sampled in the Achieving Strategy items) were emphasised in the text (Pauk, 1974).

Post-test results show an increase on Deep Motive, but quite dramatic shifts on all strategies: Surface decreased, with both Deep and Achieving increasing markedly (all these shifts from pre- to post-test were highly significant), as illustrated in Figure 7.

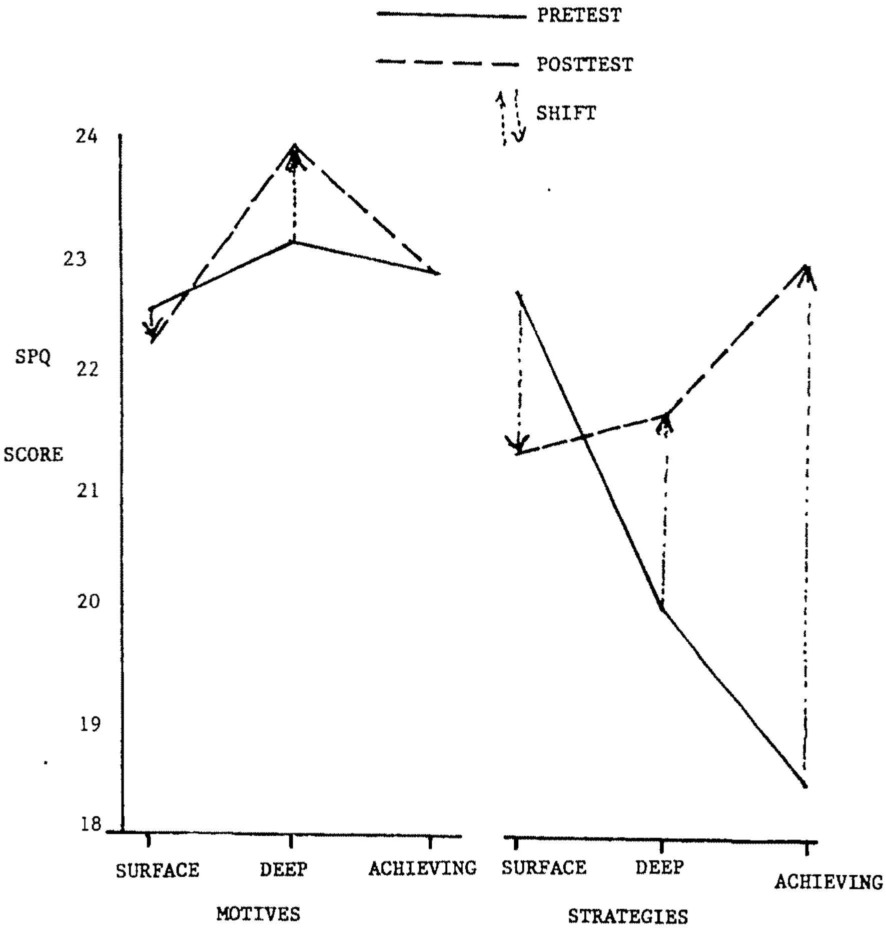
The general nature of the shift can be summarised as one moving towards better congruence with the students' motivational state. Both Deep and Achieving Strategies moved closer towards their corresponding Motives — and by comparing with Figure 6, towards the norms as well — and especially so in the case of Achieving. One distinguishing feature of the Stanford students' pretest profile was their high Achieving Motive: by the post-test, that becomes matched with a correspondingly high Achieving Strategy. Finally, it must be noted that the GPA of one of the groups increased significantly and was maintained for as long as it was monitored over the next term. Unfortunately, the corresponding data for the second group were unavailable.

In evaluating this study, it cannot be emphasised too strongly that the students were prime candidates for this kind of intervention: they were bright, very highly motivated, and themselves sought assistance. Their problem was that their strategies were out of kilter with their motives. One cannot say, therefore, that similar results would be obtained if the same procedures were used with the more typical clientele of the Student Counselling Services, whose main problem tends to be a lack of, rather than a super-abundance of, appropriate motivation.

The interest of the study in the present context is rather more theoretical than prescriptive, in that a group of students were able to improve their approaches to learning in a way that would be predicted by the congruence hypothesis, and that when deep and achieving strategies were brought into line with their motivational state, their performance improved. The mechanism for "bringing into line" appeared to involve metalearning.

FIGURE 7

STANFORD PRETEST AND POST-TEST SPQ SCORES



That is encouraging, but clearly much more research is required, in particular on the question of the relationship between the metalearning capability of the student and the nature of the intervention: that issue is alluded to in Table 10 but to date cannot be explicated in adequate detail. Another question is that of defining levels of metalearning more precisely and measuring them.

Teaching metacognitive skills, or working with metalearning capability?

In discussing the role of metacognition in education, Brown (1984) argues:

> "... some form of metacognitive theory could offer valuable contributions to the arguments about a core curriculum. Selection of problem-solving tasks in school might then be based not only upon subjects deemed to be valuable in terms of their contents, but also upon essential metacognitive skills which improved the efficiency of the children's cognitions. Metacognition may succeed where formal disciplines failed" (p. 218).

Wagner and Sternberg (1984) also suggest that metacognitive processes might be the direct target of schooling, giving the reasons that these skills are important and have a demonstrably broader degree of transfer than do cognitive strategies as such, that students are typically lacking in such skills, and that these skills are not taught already.

The logic of these views is evident, too, in terms of the present model: improving children's metacognitive skills would deepen their approaches to learning, which in turn (cf. Figure 5) would increase the structural complexity of their learning, and the amount of satisfaction derived from it.

There is, however, some difficulty with the term "metacognitive skills". Both Brown, and Wagner and Sternberg, refer to these as if they are readily definable and detachable behaviours. Viewed in that light, one might compare them to the well-known SQ3R approach to study (Robinson, 1961), which certainly may be interpreted as teaching students to be metacognitive about their study: to be aware of what it is they are to learn, and monitor both their understanding and their mastery of fact, and to remediate where necessary. Inappropriate or directive use of SQ3R, however, brings us back to the metacognitive problem: *telling* students to survey, to question, to review, etc., may become an end in itself. If students do not see the point of this, or are unmotivated to commit themselves to the deeper involvement in learning that this requires, they will treat it as just another cognitive, not a metacognitive, activity. The "metacognitive skills" simply become different, and indeed irrelevant, task content that is likely, as Wagner and Sternberg (1984) point out, to interfere with primary task content.

The present model focuses on the subject doing the metacognising, rather than on the object of metacognitive activity, such as a "skill". The problem is that while there are certain conditions under which students do seem to develop metalearning capability — such as growing older in an appropriate environment, possessing or developing an internal locus of control, or experiencing appropriate motivational states — these conditions have great momentum and their effects are usually not visible in the short term.

Thus, the intervention study carried out at Stanford might be regarded as one where cognitive strategies were taught to students who already possessed the appropriate degree of metacognitive readiness, rather than as one where metacognitive skills as such were taught, or as one where the students' metalearning capability itself had been enhanced.

There thus appear to be two quite distinct issues. First, some hard thinking needs to be done, as Brown (1984) urges, about what we precisely mean by metacognitive skills and how they may be taught. Here, metacognitive theory would define the content of teaching.

Second, there is the separate issue of metalearning capability, which raises a host of questions, the like of which are familiar from the individual differences literature: can metalearning capability be enhanced, and if so how? Is it likely to be fruitful to treat metalearning as an aptitude in aptitude-treatment interaction designs? Here, metacognitive theory would help shape the process of teaching. It is really on this second issue that the present paper has most to contribute.

Metalearning capability and learning orientation

The metalearning capability of the student is argued to determine strongly how that individual's learning process complex is structured; and following from that, the learning orientation that the individual displays. The learning process complex in fact represents an attempt to subsume many individual difference variables that are relevant to institutional learning under consistent motivational and strategic differences. The LPQ and SPQ, then, provide ways of operationalising these differences in scale and subscale profiles (see above, Table 2), which thus represent an individual's general orientation to learning: that is, a composite of motivational state and strategy deployment that is consistent over situations.

Because of the correlations between motives and strategies, there are in practice relatively few such profiles. Using '+' for 'above average' on a particular subscale (and leaving aside the question of how much above average that might be), 'O' for 'average', and '−' for 'below average', one can succinctly represent students' learning orientation profiles. Six of the most common are listed in Table 11, which strongly recall syndromes familiar to teachers and psychologists under different names, such as "independent learner", "high achiever", "under-achiever" and the like, and each of which has attracted its own favoured treatments. Table 11 also mentions briefly some of these treatments, as culled from a broad literature.

TABLE 11

SOME COMMON MOTIVE-STRATEGY PROFILES AND SUGGESTED INSTRUCTIONAL TREATMENTS

		Profile							
		Surface		Deep		Achieving			
		M	S	M	S	M	S	Description	Possible Treatment
1		+	+	0	0	0	0	Surface	High structure. clearly defined objectives, emphasise algorithms, criterion referencing, mastery learning.
	or	+	+	−	−	−	−		
2.		+	+	0	0	+	0	Surface-Achieving	Focus on organising skills and time management (Achieving Strategy) as in "Stanford Syndrome", when students very bright and achievement motivated.
	or	+	+	−	−	+	−		
3		0	0	0	0	+	+	Achieving	High structure· emphasise competition, exam technique, norm-referencing
4.		0	0	+	+	0	0	Deep	Independent study, but may need guidance towards institutional goals
5		0	0	+	+	+	+	Deep-Achieving	As for (4), perhaps higher structure.
6		0	0	0	0	−	0	Low-Achieving	Low motivation the target. attribute success to ability, failure to insufficient effort. Mastery learning to build competence and self esteem
	or	+	0	0	0	−	0		

How teachers might best use such information about student profiles involves questions of philosophy and teaching style as much as of learning theory. Logically, there are two broad options for teachers:

(1) To accept the student's orientation as given, and match instructional objectives, teaching processes, and evaluation procedures to that orientation, as in the familiar aptitude-treatment interaction (ATI) model, in order to maximise *content* learning (Biggs, 1973);

(2) To attempt to change the student's orientation where it is seen to be maladaptive in order to maximise *process* learning (Biggs, 1973). Should one, for example, teach students who are high on a surface approach in a high structure environment, with clearly defined objectives that may even be rote learned to criterion? Or should one confront such students with a low structure teaching environment which would require them to deploy what deep strategies they have in the hope of developing them further in the deep direction? Clearly, one cannot answer that question without reference both to more empirical work than is currently available,

and to value-laden issues that cannot appropriately be canvassed here: the point simply needs to be raised.

It is sufficient for present purposes, then, to draw attention to some of the ways in which the present model may help both to conceptualise and to operationalise individual differences in metacognitive capability, and to match (or mismatch) those differences with different instructional treatments. The present model does not prescribe or generate new treatments so much as to integrate a developing theoretical area — that of metacognition and in particular of metalearning — with existing teaching paradigms.

Research implications abound. Applied research might be devoted to the recognition and description of the more common motive-strategy profiles and with what classroom pathologies they appear to be associated; and, following from that, deriving through case study and intervention studies what appear to be optimal treatments for handling students with these profiles. Such treatments would include informal student-teacher interaction as well as formal ATI models in which learning orientation is matched to appropriate curricular, instructional, and/or evaluation features.

With regard to basic research, the model represented in Figure 5 could be used to derive many hypotheses. Probably the most important issues for more basic research, however, would include studies of the nature and development of metalearning, its role in mediating the connections between presage, process, and outcome, and the conditions under which metalearning capability may be enhanced.

ACKNOWLEDGMENTS. — Thanks are due to the staff and students of the participating institutions, and to the then Director of the Australian Council of Education Research, Dr. J. P. Keeves, and to Dr. Jan Lokan of ACER, for essential logistic assistance; to the Education Research and Development Committee, the Australian Research Grants Scheme, and to the University of Newcastle, for financial support; and to Mrs. Helen Ward for assistance with data processing.

REFERENCES

ANDERSON, T. H. (1979) Study skills and learning strategies. In O'NEIL, H F., and SPIELBERGER, C. D. (Eds.), *Cognitive and Affective Learning Strategies.* New York: Academic Press

APTER, M J (1982). *The Experience of Motivation.* London: Academic Press.

BIGGS, J. B. (1973). Content to process. *Aust. J. Educ.*, 17, 225-238

BIGGS, J. B. (1978) Individual and group differences in study processes. *Br. J. educ. Psychol.*, 48, 266-279

BIGGS, J. B. (1979) Individual differences in study processes and the quality of learning outcomes. *Higher Educ.*, 8, 381-394.

BIGGS, J. B. (1982). Student motivation and study strategies in university and CAE populations. *Higher Educ. Res. Dev.*, 1, 33-55.

BIGGS, J. B. (1984). Motivational patterns, learning strategies, and subjectively perceived success in secondary and tertiary students. In KIRBY, J. R (Ed.), *Cognitive Strategies and Educational Performance* New York: Academic Press.

BIGGS, J B. (in press). *Student Approaches to Learning and Studying* Hawthorn, Vic. Australian Council for Educational Research

BIGGS, J. B., and COLLIS, K. F. (1982). *Evaluating the Quality of Learning: The SOLO Taxonomy* New York. Academic Press.

BIGGS, J. B., and DAS, J. P (1973). Extreme response set, internality-externality and performance. *Br. J. soc clin. Psychol.*, 12, 199-210.

BIGGS, J. B., and KIRBY, J. R. (1983). Approaches to learning in universities and CAEs. *Vestes*, 27 (2), 3-9.

BIGGS, J. B., and KIRBY, J. R. (1984). Differentiation of learning processes within ability groups. *Educ. Psychol.*, 4, 21-39.

BIGGS, J. B., and RIHN, B. (1984). The effects of intervention on deep and surface approaches to learning. In KIRBY, J. R. (Ed.), *Cognitive Strategies and Educational Performance.* New York: Academic Press.

BROWN, A. L. (1978). Knowing when, where, and how to remember: a problem in metacognition. In GLASER, R. (Ed), *Advances in Instructional Psychology* (Vol. 1). Hillsdale, N. J.: Erlbaum.

Brown, A. L., Bransford, S. D., Ferrara, R. A., and Campione, J. C. (1983). Learning, remembering and understanding. In Mussen, P. H. (Ed.), *Handbook of Child Psychology (4th Ed.), Vol. 111, Cognitive Development*. New York: Wiley.

Brown, G. (1984) Metacognition: New insights into old problems? *Br. J. educ Stud.*, 32, 213-219.

Das, J. P., Kirby, J. R., and Jarman, R. (1979). *Simultaneous and Successive Cognitive Processes*. New York: Academic Press.

Entwistle, N. J. (in press). Motivational factors in students' approaches to learning. In Schmeck, R. (Ed.), *Learning Styles and Learning Strategies*. New York. Plenum.

Entwistle, N., Hanley, M., and Hounsell, D. (1979) Identifying distinctive approaches to studying *Higher Educ.*, 8, 365-380.

Entwistle, N, and Kozéki, B (1985) Relationships between school motivation, approaches to studying, and attainment among British and Hungarian adolescents. *Br. J. educ. Psychol.*, 55, 124-137

Flavell, J. H. (1976). Metacognitive aspects of problem solving In Resnick, L. B (Ed), *The Nature of Intelligence* Hillsdale, N J.: Lawrence Erlbaum.

Fransson, A (1977) On Qualitative difference in learning IV — Effects of intrinsic motivation and extrinsic test anxiety on process and outcome. *Br. J. educ. Psychol.*, 47, 244-257

Gammage, P. (1982). *Children and Schooling*. London: Allen and Unwin.

Gibbs, G. (1977). Can students be taught how to study? *Higher Educ Bull*, 5, 107-118.

Johnson Abercrombie, M. L. (1969). *The Anatomy of Judgement*. Harmondsworth: Penguin Books.

Keeves, J. P., and Bourke, S. F. (1976) *Australian Studies in School Performance. Vol. 1. Literacy and Numeracy in Australian Schools* Canberra Education Research and Development Committee Report No. 8.

Kirby, J. R., and Biggs, J B (1981) Learning styles, information processing abilities, and academic achievement. Final Report, Australian Research Grants Committee, Belconnen ACT.

Laurillard, D. (1979) The processes of student learning. *Higher Educ*, 8, 395-409.

Lokan, J., Boss, M. W., and Patsula, P. (1982). A study of vocational maturity during adolescence and locus of control *J vocat Behav.*, 20, 331-342

Marton, F., Hounsell, D., and Entwistle, N. (Eds) (1984) *The Experience of Learning*. Edinburgh: Scottish Academic Press.

O'Neil, M. J., and Child, D. (1984). Biggs' SPQ. a British study of its internal structure. *Br. J. educ. Psychol*, 54, 228-234.

Pauk, W. (1974). *How to Study in College*. Boston: Houghton Mifflin

Ramsden, P. (1979). Student learning and perceptions of the academic environment *Higher Educ.*, 8, 411-427.

Robinson, F. P. (1961). *Effective Study* New York Harper and Row.

Snowman, J. (in press). Learning tactics and strategies. In Phye, G. D., and Andre, T. (Eds.), *Cognitive Instructional Psychology Components of Classroom Learning*. New York· Academic Press.

Tabberer, R. (1984). Introducing study skills at 16-19. *Educ. Res*, 26, 1-6.

Taylor, E. (1984) Orientation to study: a longitudinal investigation of two degrees in one university Unpublished Ph.D. dissertation, University of Surrey

Thorndike, R. L. (1973) *Reading Comprehension Education in Fifteen Countries· An Empirical Study.* Stockholm: International Association for the Evaluation of Educational Achievement.

Wagner, R. K., and Sternberg, R. J (1984). Alternative conceptions of intelligence and their implications for education. *Rev. educ. Res.*, 54, 179-224.

Watkins, D. (1982). Academic achievement and the congruence of study motivation and strategy. *Br. J. educ. Psychol.*, 52, 260-263.

Watkins, D (1983). Assessing tertiary study processes. *Human Learning*, 2, 29-37.

Watkins, D., and Hattie, J (1981). The learning processes of Australian university students: investigations of contextual and personological factors *Br. J educ. Psychol.*, 51, 384-393.

Watkins, D., and Hattie, J (1983). A longitudinal study of the learning processes of tertiary students. Extended version of paper presented to Annual Conference, Australian Association of Research in Education, Canberra, November

(*Manuscript received* 19*th April* 1985)

Br. J. educ. Psychol., 55, 213-223, 1985

TRANSFER OF TRAINING IN THE GENERATION OF DIAGNOSTIC HYPOTHESES: THE EFFECT OF LOWERING FIDELITY OF SIMULATION

By N. C. BOREHAM
(*University of Manchester*)

SUMMARY. Practice on simulations of diagnostic problem solving can provide the knowledge base required for generating hypotheses for new problems. However, this kind of learning may be inhibited if the simulation requires the learner to practise the convergent type of hypothesis testing typical of most diagnostic situations. This prediction was tested in an experiment in which 40 adults practised on a simulation of a diagnostic task, and were then tested on their ability to generate hypotheses for a new problem. The degree of convergence/divergence of the hypothesis testing in the practice task was varied experimentally. It was found that subjects who practised the more convergent form of hypothesis testing acquired less hypothesis generating ability ($P < 0{\cdot}05$) than those who practised the more divergent form of hypothesis testing. These findings imply that, in certain circumstances, transfer performance may be improved by lowering fidelity of simulation in the practice task. This contradicts the widespread acceptance in theories of transfer that maximum similarity of the practice and criterion situations always produces maximum transfer.

INTRODUCTION

THE "fidelity problem" in the design of simulation devices is the question of how similar conditions of practice should be to the conditions in which the task is performed in order to maximise transfer of learning. In simple tasks, such as learning verbal paired associates, transfer is usually highest when the similarity of the practice and performance situations is highest (Osgood, 1949). However, the Osgood surface does not take account of the many factors which influence transfer in more complex tasks (Holding, 1976; Hammerton, 1977). A number of studies of real-world learning contradict the Osgood theory, demonstrating that transfer can actually be increased by reducing fidelity of simulation in the practice task. For instance, it has been found that if certain display features are omitted from aircraft instrument panels during practice, trainee aircrew may perform better on transfer than if they practised on the instruments in their operational condition (Grunwald, 1968; Trollip and Ortony, 1977). Such findings support the claim by Stammers and Patrick (1975) and Stammers (1981) that the conditions of simulation which maximise transfer are not necessarily those of maximum fidelity. They also pose a problem for current theories of transfer, which do not readily explain this phenomenon (Hammerton, 1977; Ward and Gow, 1982).

The present paper investigates the conditions under which a reduction in fidelity of simulation may produce an improvement in transfer performance. It is hypothesised that this will occur when both of the following conditions are met: (1) when performance depends on the amount of first task learning which occurs, e.g., when learners are switched to the criterion situation before they have all achieved mastery, and (2) when the operational conditions under which the task is performed contain factors which would inhibit the initial *learning* of the task, but not its *execution* once mastery is achieved.

The first of these conditions has received a considerable amount of research attention (Holding, 1976). Under such circumstances, there can be little doubt that superior transfer performance will be produced by the practice condition which brings about the greatest amount of learning in the time available.

A more crucial part of the hypothesis is the second of the two conditions. A familiar example of this is how learning can be speeded up by providing sources of extrinsic feedback which do not exist in the operational situation. In many cases, learning to rely on the intrinsic feedback generated by the task itself occurs more quickly by practising under such conditions. However, this is a trivial example: almost all effective training simulators are low fidelity in this respect.

Of greater theoretical significance is the question of whether transfer performance can be improved by altering the structure of the task itself. Previous researchers have investigated this possibility by making the practice task easier. For instance, if a person practises on an initial discrimination task, then transfers to a more difficult discrimination on the same dimension, performance on the latter task may be superior than if the training had been given on a task equal in difficulty to the transfer one (Sweller, 1972; Damrauer, 1976). However, superior transfer can also be produced by making the practice task more difficult, and so it is generally agreed that the concept of task difficulty is of little value for predicting transfer effects (Holding, 1976).

Greater insight into the fidelity problem may be gained by investigating the differences between the *information processing* required for *learning* a task, and that required for *performing* it. Such differences often occur in the more convergent kinds of problem solving. In such tasks, learning typically requires the deep processing of particular categories of information, whereas effective performance may require that the same information be filtered out of the information processing system altogether. Consequently, practising meeting the operational performance requirements may involve adopting information processing strategies which inhibit initial learning. This argument will be developed in the context of a specific example of problem solving.

Diagnostic problem solving

Diagnostic problem solving is important in many professions, and has received much research attention recently (for reviews, see Elstein *et al.*, 1978; Rasmussen and Rouse, 1981; Swets and Pickett, 1982; Gale and Marsden, 1983). The typical diagnostic problem arises when a system (such as the human body, an engine, or an industrial organisation) presents a symptom of malfunction, and the diagnostician has to find the cause by investigating the state of the system. Two crucial components of skill in this task are hypothesis generation (thinking up possible causes of the problem) and hypothesis testing (using data about the malfunctioning system to decide which of them are true and which are false).

"Hypothesis generation" is a blanket term covering several cognitive processes (Miroshkhina, 1969; Gettys and Fisher, 1979; Gale, 1982). One of these is the ability to state a small number (typically four to six) of the *most probable* explanations for the symptom in question (Elstein *et al.*, 1978). The present research, however, is concerned with a more fundamental process than this: the ability to generate a large (ideally, exhaustive) set of explanations for the symptom (Manning *et al.*, 1980). This is a prerequisite for the more selective type of hypothesis generation, since if a particular hypothesis cannot be generated at all it obviously can never be generated in the appropriate context.

The basic condition upon which hypothesis generation depends is the existence in the diagnostician's long-term memory of a set of explanations for the types of symptom he might encounter. This kind of knowledge base is described by Broudy (1977) as 'knowing with', and its crucial role in diagnosis and other forms of problem solving has received considerable research attention in recent years (Greeno, 1980; Feltovich, 1981).

Acquisition of ability at hypothesis generation presupposes the acquisition of this kind of knowledge base. The present study is concerned with the learning of this aspect of diagnostic problem solving through practice on simulators.

Written simulations of diagnosis

The practice environment to be investigated is that provided by written simulations of diagnosis (WSDs). Devices of this kind ("tab items") were originally developed for teaching fault-finding to maintenance technicians (Glaser *et al.*, 1954; Rimoldi, 1955). Today, they are used to provide practice in diagnosis for trainees in a broad range of professions (McGuire *et al.*, 1976; Eggen *et al.*, 1982). The basic format of the WSD comprises (*a*) a description of the system which is to be diagnosed, (*b*) a description of the symptom, and (*c*) a list of questions about the system which, if selected by the learner, provide further information about it. Typically, the learner is asked to imagine he is in the real situation and to try to diagnose the cause of the symptom without asking any redundant or irrelevant questions.

Significant learning of diagnostic skills can occur as a result of practice on this kind of device (Rimoldi and Devane, 1961; de Tornyay, 1968). In particular, it can build up the knowledge base needed for hypothesis generation, since the list of questions portrays the range of symptom-generating states that the system might have fallen into. Question lists of this kind were the basis of Suchman's (1961) method of Inquiry Training. By solving problems with them, Suchman claimed, learners acquired "operational schemata" of possible causal factors which enabled them to generate hypotheses for new problems. Evidence of this is provided in an experiment by Klein *et al.* (1969), in which exposing students to lists of possible explanations for a diagnostic problem significantly improved their scores on a test of hypothesis generation.

Rationale for predicting a negative association between fidelity of simulation and transfer performance at hypothesis generation

The application of the basic argument of this paper to the instructional context described above can now be specified as follows.

(1) The two skills of hypothesis generation and hypothesis testing are closely interdependent in reality (Gale, 1982). To adopt the training method of high fidelity simulation assumes that the learner can acquire both of them in tandem. Thus in the WSD, the question list is not only a source of data for building up a knowledge base of possible causes of malfunction but also provides information for testing hypotheses about the cause of the particular symptom observed.

(2) The typical diagnostic task is highly convergent (Rasmussen and Rouse, 1981; Swets and Pickett, 1982)—i.e., the symptom has a large number of possible causes, but only one of them (or a small number acting in combination) is actually responsible for the fault. In such circumstances, hypothesis testing involves highly selective forms of information processing. Because the actual evaluation of the hypotheses involves processes located in short-term memory (Kozielecki, 1972; Elstein *et al.*, 1978; Bouwman, 1983), the processing capacity available is severely limited. To avoid the cognitive strain which would thus be imposed by trying to judge the truth-value of every possible hypothesis simultaneously, the following selective heuristics are adopted.

(*a*) One concerns the processing of the diagnostician's long-term memory store of possible causes for symptoms. It appears that recognition of the type of symptom observed activates only a limited number of these, the bulk of the memory store never being activated at all on any one problem (Kozielecki, 1972). This kind of intuitive judgment tends to be very accurate in practice, and permits problems to be solved without actively considering every logically possible hypothesis (Johnson-Laird, 1983).

(*b*) Another concerns the diagnostician's search of the malfunctioning system for data to test the hypothesis. The typical diagnostic task environment contains much information which is irrelevant to the sub-set of hypotheses actually under test. Consequently, cognitive strain is avoided by giving selective attention only to those sources of information which are relevant to the problem. A study by Kleinmuntz (1968) found that the level of processing of such data depended on whether or not it was perceived by the subject to be relevant to deciding the truth-values of the hypotheses under test. Irrelevant information tended to be processed at a superficial cognitive level. It was not entered into long-term memory, and because of this could not subsequently be recalled. In contrast, data perceived to be relevant to the hypotheses being tested was processed at a deeper level, and could subsequently be recalled by the subject.

(3) For the purpose of acquiring the knowledge base required for hypothesis generation by practising on a WSD, the following conditions at least need to be met if the desired learning is to occur. (*a*) All the information in the list which might suggest causative factors of a kind not previously known to the learner should be assimilated into long-term memory. Furthermore, (*b*) since learning involves more than simply adding new facts to existing ones (Bruner, 1966) the existing cognitive structure itself needs to be reorganised. In the context of problem solving, Rumelhart and Norman (1978) point out that this "involves actual changes in the very categories we use for interpreting new information . . . new structures are devised for . . . imposing a new organisation on that already stored".

(4) However, these desiderata for developing the capacity of hypothesis generation may be compromised by the selective information processing heuristics adopted in convergent hypothesis testing. Not all the data in the question list which suggest causative factors will be assimilated into long-term memory if the learner selectively filters it and "dumps" items which he perceives as irrelevant to the particular symptom being diagnosed. Moreover, a general reorganisation of existing cognitive structure will not occur if the learner only activates the sub-set of schemata relevant to the individual case under consideration.

METHOD

Hypothesis

The basic hypothesis is that the convergence of the hypothesis testing typical of diagnosis results in a selective information processing strategy which pre-empts the broader processing required for acquiring a knowledge base for hypothesis generation. Consequently, in a high fidelity simulation of a convergent task, the objective of practising hypothesis testing is likely to conflict with the objective of acquiring a knowledge base from the case details.

Owing to the covert nature of the cognitive processes involved, it is not possible to test this hypothesis directly. Nevertheless, it is possible to experimentally vary the convergence of the hypothesis testing required in a simulation of a diagnostic task, and then to test whether this influences learners' ability to generate a comprehensive range of hypotheses for a new problem.

Learning materials

Students were given the objective of learning to generate a comprehensive list of possible causes for a given symptom while they practised on a simulated diagnostic task. This task was diagnosing the cause of a performance problem in an industrial organisation, "performance problem" being defined as the situation when a job is not being carried out as well as it should be. The factors which can give rise to such problems include lack of aptitude on the part of the personnel involved, insufficient training, low level of motivation, inefficient work methods and defective equipment (Mager and Pipe, 1970; Romiszowski, 1981). Careful diagnostic reasoning is needed to isolate the causal factor(s) when such problems occur (Gilbert, 1967).

Subjects

The subjects were 40 adults (34 male, 6 female) enrolled on post-experience courses which included some study of organisational behaviour and related issues in applied psychology, such as training and work motivation.

Experimental conditions

A randomised groups experiment was conducted, in which 20 of these subjects were randomly allocated to each of two practice conditions. These differed in terms of the convergence/divergence of the hypothesis testing required. The practice exercise in each condition was a short booklet containing the following:

(1) a description of an organisation (a salt packaging plant);

(2) a brief description of the performance problem (contamination of the product);

(3) a list of 22 questions about various aspects of the organisation and its personnel (Appendix 1). All the above materials were based on a case study of a real-world event.

(4) Yes/no answers to the questions, each of which was keyed by reference to the case material to provide the information needed to solve the problem.

Students were assigned hypotheses and instructed to test them by reference to the information in the question list. The difference between the conditions was that in the more convergent Condition 1, students were assigned just one hypothesis, while in the more divergent Condition 2 they were assigned five. The specific instructions were: "It has been suggested that this performance problem was caused by" *Condition 1:* "The following factor"; *Condition 2:* "One (or a combination of) the following five factors." This was followed by the hypothesis(es) to be tested, and the instruction: "Each of the following questions refers to this organisation and its performance problem. Find out by consulting these questions whether the performance problem was caused by (any of) the above factor(s)." The instructions were accompanied by a true/false response format printed alongside the hypotheses that were to be tested.

Five hypotheses were used: (1) "The inadequacy of the equipment"; (2) "the workers' lack of basic aptitude for the job"; (3) "a low level of motivation among the workers"; (4) "poor training"; (5) "inadequate work methods". Each student in Condition 2 was given all five of these to test. Each student in Condition 1 was given one to test, four students in this condition being randomly allocated to each of the five hypotheses to equalise the content of the training materials between groups.

In informational terms, the difference between the two conditions lay in the number of items in the question list which provided information relevant to the assigned task. Most of the items were specific to individual hypotheses. The result was that, in Condition 1, only a minority of the questions were relevant to the truth-value of the single assigned hypothesis, whereas in Condition 2 every question was relevant to the truth value of some of the five hypotheses at least.

This was assessed by two judges (both occupational psychologists) who were asked to decide whether each question was "relevant" to each of the five hypotheses (i.e., provided information which either increased or decreased the likelihood of that factor being the cause of the performance problem), or "irrelevant" (i.e., did not alter the prior probability of the hypothesis being true). The raters carried out this task independently. They agreed on 95 of the 110 hypothesis-question pairings (86.4 per cent). The detailed ratings revealed that in Condition 1 only about a quarter of the items in the question list provided relevant information, whereas in Condition 2 all the items provided relevant information. Averaging the two sets of ratings, it was found that of the 22 questions, 4·0 were judged relevant to hypothesis 1; 5·0 to hypothesis 2; 7·0 to hypothesis 3; 6·5 to hypothesis 4; and 4·0 to hypothesis 5. Thus, in Condition 1, the student was faced with a list of 22 questions, of which on average only 5·3 provided information relevant to the truth-value of the assigned hypothesis. However, each of the 22 questions was judged by both raters to be relevant to the truth-value of at least one of the five hypotheses. Thus in Condition 2, *all* the items were relevant to the truth-value of the assigned hypotheses.

Procedures

The experiment was conducted in a normal classroom setting. The subjects were taught and tested in small groups in sessions lasting approximately one hour. The procedure was as follows. (1) An explanation of the concept of a performance problem was given, it being stressed that such problems could be caused by a wide variety of factors. (2) The aim of the session was described as that of developing each participant's ability to think up as wide a variety of possible causes of a given performance problem as he or she could. (3) As an example of the task they were being trained for, the students were given an untimed pre-test of the ability to generate hypotheses (Appendix 2). (4) The practice exercises were administered, students being reminded that the aim was to familiarise them with the broad range of possible causes for performance problems which were modelled in the question list. They were also told that they would be tested afterwards on an instrument similar to the pre-test. Students were given 15 minutes to work through the booklets. Everyone, in fact, completed the hypothesis testing within that period, retaining the booklets to study until the expiry of the time limit. (5) An untimed post-test of hypothesis generation was administered, following the procedure for the pre-test (Appendix 2).

RESULTS

The pre- and post-test scripts were scored for the number of different categories of hypothesis which they contained. A further team of three raters were recruited to undertake the scoring, all of whom were experienced occupational psychologists. Five categories of hypothesis (aptitude, equipment, training, motivation and work methods) were defined, each category corresponding to one of the five hypotheses used in the practice task. The raters were asked to indicate which of these five categories were mentioned in each candidate's script. They were given suggested definitions of the categories for the pre-test, but were left to form their own definitions for the post-test. Each rater scored the pre-test first, then the post-test. Scripts were presented in random order, and the raters knew nothing about the existence of different experimental conditions. Inter-rater reliabilities were high: 0·88, 0·80 and 0·83 for the pre-test scoring and 0·77, 0·80 and 0·78 respectively for the post-test scoring.

Each student was awarded an aggregate score for the number of categories mentioned in his list by averaging the scores awarded by the three raters. A difference-score was then calculated for each student, representing the increase in the number of categories he mentioned in the post-test over the number he mentioned in the pre-test. Since the two tests were based on different material, no meaning can be attached to the absolute size of these scores. However, the difference between the experimental group means does indicate the effect of varying the convergence of the hypothesis testing. As Table 1 shows, students in the more divergent Condition 2 performed better as a result of practice than students in the more convergent Condition 1.

An analysis of the regression of the post-test scores on the pre-test scores within each group indicated that the regression lines were non-parallel, and that the regression

TABLE 1

PRE- AND POST-TEST SCORES FOR THE NUMBER OF CATEGORIES OF HYPOTHESIS GENERATED

	Pre-test	Post-test	Difference
Condition 1 mean	3·10	3·17	0·07
(SD)	(0·94)	(0·87)	(1·13)
Condition 2 mean	2·58	3·45	0·87
(SD)	(1·33)	(0·99)	(1·10)

slopes were substantially different from unity (the intercepts and slopes respectively being 2·55 and 0·20 (Condition 1) and 2·32 and 0·44 (Condition 2)). This made it impossible to perform either an analysis of variance on the difference-scores, or an analysis of covariance on the post-test scores with the pre-test scores as the covariant (Edwards, 1968). The significance of the difference between the conditions was therefore estimated by non-parametric methods.

The students were categorised into two groups: those whose scores had improved as a result of practice, and those whose scores had not improved. They were cross-classified by experimental condition, and a two-tailed Fisher's exact test was computed. This showed a significant difference ($P<0\cdot05$) in favour of Condition 2.

TABLE 2

NUMBER OF STUDENTS WHOSE POST-TEST SCORE SHOWED AN IMPROVEMENT ON THEIR PRE-TEST SCORE

Change from Pre-test to Post-test	Condition 1	Condition 2
Improvement	9	16
No Improvement	11	4

Fisher's exact test (two-tailed) $P<0\cdot05$

DISCUSSION

This study demonstrates that the more convergent hypothesis testing is, the more it will tend to interfere with acquisition of the ability to generate hypotheses during practice on this kind of simulation device.

Several explanations are possible for the transfer effect that was observed. (1) The results are broadly consistent with recent findings about context effects in the free recall of narrative prose (di Sibio, 1982). When the same prose passage is read by randomised groups of subjects under different thematic titles, different groups tend to recall different details from the passage. The factor which determines whether or not an item is recalled is its relevance to the context suggested by the title (Anderson *et al.*, 1978). On this view, the convergent and divergent hypothesis testing tasks in the present study would have activated different cognitive sets, the former narrower than the latter. Consequently, students in Condition 1 would have assimilated less information from the question list into long-term memory and would have been less able to generate a comprehensive range of hypotheses on the post-test. This explanation presupposes the cognitive process identified by Kleinmuntz (1968), whereby information which is irrelevant to the hypothesis under test is not subjected to deep processing, and thus never enters long-term memory.

(2) An alternative explanation is based on the assumption that test performance reflected recall of material acquired before this experiment was conducted. The experience of testing a particular hypothesis might have caused retroactive inhibition of any pre-existing knowledge that was irrelevant to that hypothesis, resulting in students in Condition 1 recalling less from memory than those in Condition 2. Evidence of this kind of inhibition in diagnosis is cited by Kozielecki (1972). Early in the diagnostic act, he found, problem solvers were able to recall from memory a broad range of possible hypotheses for a given symptom. However, once the process of testing a sub-set of hypotheses commenced, the recall of additional hypotheses from long-term memory was inhibited.

(3) Another explanation must be considered. Since the experimental conditions represented real-world tasks, several factors were inevitably confounded in the treatment differences. Besides differing in terms of the number of items in the question list which were relevant to the hypothesis testing task, Conditions 1 and 2 also differed in the *amount* of hypothesis testing required. The results may be attributable to this alone. However, even if this were the sole explanation for the findings, it would not affect the conclusion that the more convergent the hypothesis testing, the more it tends to interfere with acquisition of the knowledge base required for hypothesis generation. For to vary the convergence of a hypothesis testing task inevitably involves varying the number of hypotheses tested, provided that the content of the task remains unchanged.

More research is needed to decide between these explanations for the results that were obtained. Nevertheless, the results do have significant implications, both for educational practice and for the theory of transfer.

Implications for practice

The chief practical implication is that if learners are transferred to the criterion situation before all of them have achieved mastery (a common occurrence in most education and training), high fidelity of simulation in a WSD is not the optimal condition of practice, in so far as the criterion task is a convergent one. Such simulations might be improved by introducing constraints which make students do more of the information processing necessary for learning than the operational task itself evidently does. Thus, to develop ability at hypothesis generation, it would be desirable to make learners test a wider range of hypotheses than may actually be necessary for solving the typical diagnostic problem. Under the conditions represented here, this implies that transfer performance could be improved by reducing fidelity of simulation.

The results of this study point the need for change in some current training methodologies. For instance, in medical education it is common for students practising diagnosis to be told to imitate the highly selective information processing strategies of the expert physicians whom they accompany on ward rounds (Cutler, 1979). However, the present results suggest that such advice would retard acquisition of a knowledge base of aetiological factors, by restricting familiarisation with the wide range of data available about individual patients. An item of information which happens to be irrelevant for diagnosing one case may nevertheless be very relevant for diagnosing others. Although the expert physician (who already possesses a rich memory store of possible causes of disease) can afford to discard such information in the interests of diagnostic efficiency, that information might be vital for building up the knowledge base needed by the student.

Of course, there is more to diagnostic expertise than the possession of a rich knowledge base of aetiology. The real-life task of diagnosis requires hypotheses to be generated spontaneously, and in the case of the expert these generally comprise no more than four or five of the most likely explanations for the symptom at hand (Elstein *et al.*, 1978). Practice on WSDs (which only requires choosing among *given* hypotheses) might not be sufficient to develop such skill, and supplementary forms of training might then be necessary. Nevertheless, it would be a prerequisite for developing the art of selective hypothesis generation that the learners already possessed a full knowledge base of possible causal factors. And as we have shown, high fidelity of simulation would not provide the optimal practice conditions for achieving this.

Implications for theory

Current theories of transfer are aptly described by Fotheringhame (1984) as attempts to impose order on a very piecemeal range of experiments. The behaviour studied in transfer research has tended to be performance on simple S-R laboratory tasks which in general have been unrepresentative of the kinds of activity which interest teachers and industrial trainers. This has resulted in a widespread acceptance that maximum fidelity always produces maximum transfer. However, in view of contradictory evidence from the world of occupational training, there is a need to develop new formulations of transfer theory which can accommodate such data.

The present paper has described conditions under which lowering fidelity of simulation will improve transfer. The crucial condition is when the performance requirements inhibit the learning of particular task components, but not their execution once mastery is achieved. This might occur, for instance, when deep processing is required for learning a task, but is inhibited by the information processing limitations imposed by the performance requirements. Thus the theoretical implication of the present research is that maximum transfer should be predicted on the basis of whichever practice conditions maximise learning of the skills required for criterion performance, irrespective of their similarity to the transfer situation. In other words, transfer theory should be about learning instead of about task similarity.

Acknowledgments. — Thanks are due to Mr. M. H. Toye, Dr. R. B. Stammers and Professor James Rushton for constructive comments while this research was in progress.

REFERENCES

Anderson, R. C., Spiro, R. J., and Anderson, M. C. (1978). Schemata as scaffolding for the representation of information in connected discourse. *Am. educ. Res. J.*, 15, 433-440.

Bouwman, M. J. (1983). Human diagnostic reasoning by computer. an illustration from financial diagnosis *Management Sci*, 29, 653-672.

Broudy, H. S. (1977). Types of knowledge and the purposes of education. In Anderson, R. C., Spiro, R. J., and Montague, W E (Eds.), *Schooling and the Acquisition of Knowledge*. New York: Wiley

Bruner, J S (1966). *Toward a Theory of Instruction*. Cambridge, Mass.: Harvard University Press

Cutler, P. (1979) *Problem Solving in Clinical Medicine from Data to Diagnosis*. Baltimore: Williams and Wilkins Co.

Damrauer, B. R. (1976). Conceptual level sorting in young children: grouping on the basis of class membership in the presence of stimulus competition. *Percept. mot. Skills*, 43, 19-24.

de Tornyay, R. (1968). Measuring problem solving skills by means of the simulated clinical nursing problem test. *J. nursing Educ.*, 7, 3-8 and 34-35.

di Sibio, M. (1982) Memory for connected discourse. a constructivist view. *Rev. educ. Res.*, 52, 149-174.

Edwards, A. L. (1968). *Experimental Design in Psychological Research (3rd edit.)*. London: Holt, Rinehart and Winston.

Eggen, P. D., Kauchak, D. P., and Harder, R. J. (1982) *Strategies for Teachers: Information Processing Models in the Classroom* Englewood Cliffs, N.J.: Prentice-Hall.

Elstein, A. S, Shulman, L. S., and Sprafka, S. A. (1978) *Medical Problem Solving*. Cambridge, Mass.: Harvard University Press.

Feltovich, P. J. (1981). *Knowledge Based Components of Expertise in Medical Diagnosis*. University of Pittsburg, Learning Research and Development Center: Technical Report No. PDS-2

FOTHERINGHAME, J. (1984). Transfer of training: a field investigation of youth training. *J. occup. Psychol.*, 57, 239-248.

GALE, J. (1982). Some cognitive components of the diagnostic thinking process *Br. J. educ. Psychol.*, 52, 64-76.

GALE, J., and MARSDEN, P. (1983). *Medical Diagnosis* Oxford: Oxford University Press.

GETTYS, C., and FISHER, S. D. (1979) Hypothesis plausibility and hypothesis generation. *Organiz. Behav. hum. Performance*, 24, 93-110.

GILBERT, T. F. (1967). Praxeonomy: a systematic approach to identifying training needs. *Manag. Personnel Q.*, 6, 20-33.

GLASER, R., DAMRIN, D. E., and GARDNER, F. M (1954). The tab item: a technique for the measurement of proficiency in diagnostic problem solving tasks. *Educ. psychol. Measur.*, 14, 283-293.

GREENO, J. G. (1980). Trends in the theory of knowledge for problem solving. In TUMA, D T., and REIF, F. (Eds.), *Problem Solving and Education. Issues in Teaching and Research.* Hillsdale, N.J.: Lawrence Erlbaum.

GRUNWALD, W. (1968). An Investigation of the Effectiveness of Training Devices with Varying Degrees of Fidelity. Ph.D. Thesis, University of Oklahoma.

HAMMERTON, M. (1977). Transfer and simulation. In INSTITUTE OF MEASUREMENT AND CONTROL, *Human Operators and Simulation.* London: Institute of Measurement and Control.

HOLDING, D. H. (1976). An approximate transfer surface. *J. motor Behav.*, 8, 1-9.

JOHNSON-LAIRD, P. N. (1983). Thinking as a skill. In EVANS, J. St. B.T., *Thinking and Reasoning*. London: Routledge and Kegan Paul.

KLEIN, S P., FREDERIKSEN, N., and EVANS, F. R. (1969). Anxiety and learning to formulate hypotheses *J. educ. Psychol.*, 60, 465-475.

KLEINMUNTZ, B (1968). The processing of clinical information by man and machine In KLEINMUNTZ, B. (Ed.), *Formal Representation of Human Judgement.* New York: Wiley.

KOZIELECKI, J. (1972). A model for diagnostic problem solving. *Acta Psychol.*, 36, 370-380.

MAGER, R. F., and PIPE, P. (1970). *Analyzing Performance Problems.* Belmont, Calif.. Fearon Publishers.

MANNING, C., GETTYS, C., NICEWANDER, A., FISHER, S., and MEHLE, T. (1980). Predicting individual differences in generation of hypotheses. *Psychol. Rep.*, 47, 1199-1214.

MCGUIRE, C (1976). Simulation technique in the teaching and testing of problem solving skills. *J. Res. sci. Teaching*, 13, 89-100.

MCGUIRE, C H., SOLOMON, L. M., and BASHOOK, P. G. (1976). *Construction and Use of Written Simulations.* New York: The Psychological Corporation, Harcourt Brace Jovanovich.

MIROSHKHINA, E. A. (1969). On the relation between the structure of a problem and the structure of heuristic search in man. In PUSHKIN, V. N. (Ed.), *Problemy evristiki* (Problems of Heuristics) Moscow Academy of Pedagogical Sciences of the U.S.S.R.

OSGOOD, C. E. (1949). The similarity paradox in human learning: a resolution. *Psychol. Rev.*, 56, 132-143.

RASMUSSEN, J., and ROUSE, W. B (Eds.) (1981). *Human Detection and Diagnosis of System Failures.* New York: Plenum.

RIMOLDI, H. J. A. (1955). A technique for the study of problem solving *Educ. psychol. Measur.*, 15, 450-461.

RIMOLDI, H. J. A., and DEVANE, J. R. (1961). *Training in Problem Solving.* Chicago: Loyola University Psychometric Laboratory (Mimeo).

ROMISZOWSKI, A. J. (1981). *Designing Instructional Systems.* London: Kogan Page.

RUMELHART, D. E. and NORMAN, D. A. (1978). Accretion, tuning and restructuring: three modes of learning. In COTTON, J. W., and KLATZKY, R. L. (Eds.), *Semantic Factors in Cognition.* Hillsdale, N.J.: Erlbaum.

STAMMERS, R. B. (1981). Theory and practice in the design of training simulators *Programmed Learn. educ. Technol.*, 18, 67-71.

STAMMERS, R., and PATRICK, J. (1975). *The Psychology of Training.* London: Methuen.

SUCHMAN, J. R. (1961). Inquiry training: building skills for autonomous discovery. *Merrill-Palmer Q.*, 7, 147-169.

SWELLER, J. (1972). A test between the selective attention and stimulus generalization interpretations of the easy-to-hard effect. *Q. Jl. exp. Psychol.*, 24, 352-355.

SWETS, J. A , and PICKETT, R. M. (1982). *The Evaluation of Diagnostic Systems.* New York: Academic Press.

TROLLIP, S., and ORTONY, A. (1977). Real-time simulation in computer-assisted instruction. *Instruct. Sci.*, 6, 135-149.

WARD, J., and GOW, L (1982). Programming generalization: a central problem area in educational psychology. *Educ. Psychol.*, 2, 231-248.

(*Manuscript received* 12*th February*, 1985)

APPENDIX I

The list of questions

Is the filling machinery malfunctioning? No.
Do the baggers have good enough eye-hand co-ordination? Yes.
Are the baggers distracted by the foreman's frequent requests for production figures? No.
Does the bagging plant need to be re-designed? No.
Was the training period too short? No
Do the baggers have sufficient stamina to maintain a high standard of production? Yes
Does the speed at which orders have to be packed make careful handling impossible? No.
Are the baggers equipped with the most hygienic filling machinery available? Yes.
Do the baggers have all the skills and knowledge to do their job correctly? Yes.
Is the training method (sitting beside Nellie) capable of imparting sufficient skill? Yes.
Do the baggers feel they are receiving adequate rewards for their work? No.
Have the baggers been told to check the quality of their output? Yes.
Do the baggers need a course of refresher training? No
Do the baggers worry that more efficiency would lead to redundancies? No.
Did the baggers lack the intelligence to learn the correct packing procedure? No
Could the baggers package salt hygienically if their lives depended on it? Yes.
Is there a shortage of replacement air filters? No.
Was hygiene stressed in the training programme? Yes.
Are the baggers conscientious about their work? No.
Would any semi-automatic bagging process like this one yield satisfactory results? Yes
Have the baggers become bored with their work? No.
Did the people who were recruited for the job of bagger have all the prerequisite aptitudes? Yes.

APPENDIX 2

The Pre-Test

1 Read the following description of a performance problem:

The Organisation
This was an open-cast mine in equatorial Africa. The open-cast method of mining is to remove the top layer of soil and rock, excavate the ore, then dump the top layer back into the cavity. Thus teams of men were at work in this mine blasting and excavating, while giant tracked vehicles continually moved among them transporting the soil, rock and ore

The Performance Problem
There were frequent accidents, in which workmen were run over by the vehicles.

2 List as many possible explanations for this problem as you can, in the space below·

The Post-Test

1. Read the following description of a performance problem.

The Organisation
This was a machine-tool manufacturing company in Korea, employing 150 skilled engineering craftsmen.

Performance Problem
The quality of the machine tool parts being made was very poor, and consequently the company was losing business to foreign imports

2 List as many possible explanations for this problem as you can, in the space below·

Br. J. educ. Psychol., 55, 224-232, 1985

GENDER SALIENCE AND SOCIAL IDENTITY: THE IMPACT OF SEX OF SIBLINGS ON EDUCATIONAL AND OCCUPATIONAL ASPIRATIONS

By DOMINIC ABRAMS, KAREN SPARKES, and MICHAEL A. HOGG*
(Department of Psychology, University of Dundee)

Summary. It is proposed that an understanding of sex differences in educational and occupational aspirations of adolescents would benefit from a perspective which stresses gender identification. The social identification approach (Turner, 1982) hypothesises that these differences are a function of the *degree* of gender identification, which in turn is a function of the salience of sex. According to McGuire and McGuire's (1981) distinctiveness postulate, gender becomes more salient as the number of opposite sex siblings increases. In order to explore these gender salience effects 81 male and female sixth form pupils completed a questionnaire concerning their self-perceptions, examination performance, occupational choice, and comparisons with members of each sex. Large sex differences were obtained, showing that educational and occupational ambitions were inhibited among females. Correlational analyses revealed that number of opposite sex siblings was more closely related to responses than were number of same sex, or total number of, siblings. In particular, females' choice of science O-levels decreased as number of brothers increased. The results are taken to suggest that sex of sibling effects have an intergroup dimension which should be the focus of further research.

INTRODUCTION

It is well known that the choice of subjects taken at school is largely determined by the child's sex. Boys tend to opt for science based subjects whereas girls concentrate their efforts on arts subjects (Sharpe, 1976; Sutherland, 1981).[1]

While this trend can be partly explained by sex role socialisation and the dominance of traditional sex-role ideology (Del Boca and Ashmore, 1981), differential achievement motivation (Alper, 1973) and 'fear of success' (Horner, 1972; but see Weinreich-Haste, 1984), there is still room for a more *social* psychological level of analysis — one that stresses identity (cf. Sherif, 1982, 1984). Abrams and Condor (1984), in taking a 'social identification' approach (Tajfel and Turner, 1979), have suggested that gender identity is partly constructed by *intergroup* relations between the sexes. Tajfel and Turner's social identity theory posits that part of the self-concept is made up of identifications with various groups and social categories. It is further assumed that individuals maintain a positive self-concept by making social comparisons with others in such a way that self receives a more favourable evaluation than other. When social identity is salient these self-other comparisons are made on the basis of group membership. People desire to make their own group (in-group) positively distinctive in relation to other groups (out-groups). Abrams and Condor (1984) suggest that one mechanism by which the roles of members of the two sexes remain

---

* Michael Hogg is now at Macquarie University, Sydney.

[1] DES figures for GCE O-level and A-level qualifications gained in the years 1980–81, and 1982–83 confirm this point. In 1982–83 the percentages of males and females gaining O-level English qualifications were 34% and 46%, while those gaining Mathematics qualifications were 34% and 29% respectively. 22% of males and 9% of females obtained O-level passes in Physics In contrast 11% of males and 20% of females passed O-level French. This pattern is repeated at 6th form level. Of those obtaining 2 or more A-level passes in 1982–83 38% of males and 17% of females had taken only Science subjects, 30% of males and 49% of females had taken only Arts and Social Studies subjects while 32% of males and 33% of females had taken a mixture of Science and Arts subjects.

distinct is that, at an intergroup level, each sex strives to maintain its positive distinctiveness from the other (see Tajfel and Turner, 1979). The pervasiveness of *traditional* sex role ideology leads females to seek positive distinctiveness by adhering to traditionally feminine role demands whereas males seek it by adhering to traditionally masculine role demands (cf. Williams and Giles, 1978).

Abrams and Condor (1984) proposed that the consequences of self-categorisation in terms of (and identification with) one's gender are different for each sex and at different ages. In the course of adolescence, the criterial stereotype attributes (see Turner, 1982) of masculinity are increasingly *congruent* both with the values of the education system (Blackstone, 1976) and with positive social evaluation. Identification with gender for males is therefore relatively unproblematic. In contrast, among female adolescents, the criterial stereotypic attributes of femininity are increasingly *discrepant* from those associated with achievement within the education system. Females may come to feel increasingly excluded from high status occupations and to devalue their academic achievements (cf. Katz, 1979; Sarah *et al.*, 1980). One consequence of this is that many females may preserve a positive social identity *only* by withdrawing from potentially negative intergroup comparisons with males in the field of education and occupations.[2] Furthermore, if the intergroup comparison is more problematic for females than for males it seems likely that variations in the *salience* of their gender will have more serious educational ramifications than for males.[3]

Distinctiveness as a source of gender salience

One factor which supposedly raises the salience of gender is 'solo status' or numerical distinctiveness (Taylor *et al.*, 1978; Taylor, 1981). To be the only *girl* in a family of four children will make gender identity more salient for that girl than if all of her siblings are also female. McGuire and McGuire (1981, 1982, 1984) have demonstrated that gender is indeed a more salient part of identity among children from households in which their own gender was in the minority. Their theory, however, is based on an overly *cognitive* analysis of identity — one which we do not entirely share (see also Tajfel, 1984). If the cognitive explanation were adequate, there should be no sex differences in identity salience in response to minority status. However, McGuire *et al.* (1979) found that, even with equal numbers of each sex in the household, females' gender identity was more salient than males'. Females seemed to have a lower threshold beyond which the presence of the opposite sex served to raise the salience of their own gender (see also Leventhal, 1970).

Sex of sibling effects

Much earlier, Koch's (1955) study of siblings found that females with male siblings were more competitive and tenacious than those with female siblings, while males were not differentially affected by the sex of their siblings. Koch suggested that this was because males' greater competitiveness made them less susceptible than females to modelling effects. Sutton-Smith and Rosenberg (1970) found that the effects of number of opposite sex siblings also included 'counteractive phenomena', especially among boys with two sisters, who tended to exhibit more masculine traits than those with only one sister. Similarly, Rosenberg and Sutton-Smith'(1968) found that fathers' masculinity was inversely related to the number of males, but positively related to the number of females, within the family. Sutton-Smith and Rosenberg (1970) went on to infer that "males with 'too much feminine' influence may indeed have been resisting or counteracting it" (p. 35). Thus it appears that there is some justification for assuming a link between gender salience and sex-role differentiation, which may arise from relationships within the family. In addition, the general consensus that the most important psychological effects of siblings are largely restricted to childhood (Sutton-Smith and Rosenberg, 1970) may be attributable to the fact that young children are the focus of some of the most interesting research in this area (e.g. Dunn and Kendrick, 1982; Dunn, 1984).

---

[2] This effect seems to occur during adolescence and at secondary school, but not so much during primary school years. Thus, the occupational and educational trajectories of males and females diverge most clearly in late adolescence. While debate exists over how far the education system and curricula are responsible for this (see Sutherland, 1981, Walker and Burton, 1983) we wish to emphasise the role of the increasing salience of gender identity in providing_a focus for self-evaluation and behavioural self-regulation.

[3] It has been argued that working class females are dealt a double blow since both their class and their gender inhibit their aspirations (cf. King, 1971, Purvis, 1981, Anyon, 1983).

A large proportion of correlational research on siblings has been concerned with issues such as birth order (e.g., Belmont and Marolla, 1973; Zajonc and Markus, 1975) or with the effects of a number of siblings irrespective of their sex (Terhune, 1974). The general conclusion that smaller families are associated with higher academic and occupational achievement (Marjoribanks, 1976) is not usually qualified by reference to sex composition of siblings. The review by Wagner *et al.*, (1979) of 'sibling-constellation effects' devotes only one paragraph (out of 56 pages) to sex of sibling, and that is in a subsection on Parent–Child Relations.

The preceding discussion reveals both a theoretical question and an empirical gap in the literature. Theoretically it is important to understand how gender salience actually affects the self-evaluations, aspirations, and eventual life-chances of adolescents as they pass through their school years. Empirically there is a need to demonstrate the existence of a link between an hypothesised cause of gender salience and an array of psychological and behavioural indices of gender salience.

On the basis of the McGuires' work it is predicted that the salience of gender identity is positively related to the number of opposite sex siblings. A social identification approach suggests that gender salience will encourage intergroup comparisons with the opposite sex, increased self-stereotyping (Turner, 1981, 1982) in terms of gender, and hence sex-typed behaviour and self-evaluations. Consequently, the relatively neutral cognitive *salience* of gender becomes imbued with different meanings and evaluative connotations for each sex depending on the social context and pervasive sex role ideology (Condor and Abrams, 1984).

A second prediction then, is that variations in gender salience will have a greater impact on the academic and occupational aspirations of females than on males. This is because traditional male gender identity is more compatible with such aspirations than is female gender identity.

The present study employed a 2 (sex) × 2 (school year) design in which a questionnaire was administered to measure: academic achievement; educational and occupational aspirations; self-perceptions; and intergroup comparisons. (School year is included as a factor simply in order to account for variation in the amount of education received.) Information concerning number and sex of siblings was also collected for the purpose of correlational analyses.

METHOD

Subjects

Fifty male and 31 female 16- to 18-year-old pupils comprising the entire upper and lower sixth forms of a large city comprehensive school in South Wales served as subjects. Their socio-economic background ranged from working class to lower middle class, and all were registered on GCE A-level courses.

Procedure

A 25-item questionnaire was administered to the pupils, in mixed sex groups of 12 to 15, during the morning tutorial period. A female experimenter presented the study as one concerning "the proportion of pupils taking arts and science subjects". In fact, the questionnaire contained a variety of questions. The first page asked subjects to use 7-point scales (1 = strongly disagree, 7 = strongly agree) to respond to seven questions. These asked subjects to indicate: how intelligent and sociable they thought themselves; the extent to which they compared their school performances mainly with that of girls or of boys; whether they thought girls worked harder than boys, or vice versa, or both worked equally hard.

On the next page, subjects were asked to list the GCE O-level subjects they had taken, and the grades obtained, and the GCE A-level subjects they were studying and the grades they expected to obtain in those examinations. From these data it was possible to ascertain the number of science and arts subjects taken by each student and the average number of grades points obtained or expected.

The final set of questions was open-ended. Pupils were asked to write down their ideal and realistic occupational choices, and how many brothers, and how many sisters they had. The occupational choices were coded in terms of socioeconomic status (1 = highest, 5 = lowest, according to the Registrar General's (1980) classification of occupations).[4]

RESULTS

The seven-point scale data, status of occupational choices, and examination data were analysed by 2-way (Sex × School Year) ANOVAs. There were no significant effects of school year, but several significant main effects of sex (see Table 1).

Males took more science and fewer arts O-levels than did females. However, neither previous nor expected performance differed. There were no reliable differences in either the overall number of examinations entered for or overall previous or expected grades.

Males were more likely to regard themselves as being of above average intelligence than were females. They also claimed that boys work harder than girls and they compared themselves exclusively with other boys more than did females. Finally, males expected to enter into higher status occupations more than did females although there was no difference in the status of *ideal* occupation choices.

Siblings

Total number of siblings, number of brothers and number of sisters were correlated with responses to the dependent measures (see Table 2). For total number of

---

[4] The final question asked: 'Do you think, in a marriage, that it should be the husband who works and the wife who stays at home and looks after any children? Why/Why not?' On the basis of their answers to this question subjects were classified as holding either traditional (answering 'yes') or non-traditional (answering 'no') sex role orientations. This variable has been found to be an important determinant of attitudes towards the sexes in previous research (Condor and Abrams, 1984) and it was thought wise to include it as a factor in a precautionary analysis in the present study.

The precautionary analysis with respect to sex-role orientation (traditional *versus* non-traditional; Ns = 37 and 44, respectively) revealed few systematic findings and these had no direct relevance to the hypotheses under examination.

T-tests revealed that only two measures were affected significantly. Traditionals were less likely to compare their school performance with that of other girls ($\bar{x} = 2.38$) than were non-traditionals ($\bar{x} = 3.30$; $t(79) = 2.39$, $P < 0.05$). Traditionals were also less convinced that girls and boys work equally hard ($\bar{x} = 4.00$) than were non-traditionals ($\bar{x} = 5.11$; $t(79) = 2.78$, $P < 0.01$). A chi-squared analysis of the distribution of male and female subjects in the two sex-role orientation categories revealed that while males were evenly divided between non-traditional (28) and traditional (22), females clearly favoured the former (22) over the latter (9) roles ($\chi^2 = 5.61$, $P < 0.05$). Despite this confounding of sex and sex-role orientation the significant effects of role orientation were not simply artefacts of differential distribution of males and females in the traditional and non-traditional groups. None of the dependent measures which revealed significant effects of sex showed any sign of being influenced by sex-role orientation, and vice versa.

TABLE 1

MAIN EFFECTS OF SEX (df = 1,79)

Item	Males	Females	F
1. Regard self as sociable	5·46	5·44	0·13
2 Own IQ is above average	5·40	4·58	5·93*
3. Compare school performance with boys'	4·55	2·71	20·79***
4. Compare school performance with girls'	2·56	3·39	2·49
5. Boys work harder than girls	4·10	2·90	8·29**
6. Girls work harder than boys	3·22	3·00	0·28
7. Both work equally hard	4·40	4·94	2·14
8. Realistic occupational choice	1·46	2·06	8·16**
9. Ideal occupational choice	1·66	1·84	0·74
10. Number of GCE O-levels taken	7·80	8·45	2·95
11 Average O-level grade	1·72	1·55	0·53
12. Number of Arts O-levels taken	2·74	4·81	36·25***
13. Average Arts O-level grade	1·55	1·55	0·01
14. Number of Science O-levels taken	5·06	3·64	17·69***
15. Average Science O-level grade	1·78	1·51	1·21
16. Number of GCE A-levels taken	2·72	2·64	0·06
17. Average expected A-level grade	2·06	1·82	0·73
18. Number of Arts A-levels taken	0·46	1·45	15·37***
19. Average expected Arts A-level grade	1·10	2·04	5·18*
20. Number of Science A-levels taken	2·26	1·19	10·76**
21. Average expected Science A-level grade	3·03	1·59	13·26***

* P<0·05; ** P<0·01; *** P<0·001.

Items 1 to 7 are scored as seven-point scales (1 = strongly disagree, 7 = strongly agree). Items 8 and 9 were scored as a five-point scale (1 = highest status, 5 = lowest status). Items 11, 13 and 15 ranged from 3 (highest pass) to 1 (lowest pass) and 0 (fail). Items 17, 19 and 21 ranged from 5 (highest pass) to 1 (lowest pass) and 0 (fail).

siblings (irrespective of their sex) there were only two significant correlations. Subjects with more siblings were less likely to have taken arts O-levels, a relationship which remained significant among male but not among female subjects.

As predicted, the lack of effects of total number of siblings was partly due to the different effects arising from numbers of siblings of each sex. As number of brothers

TABLE 2

SIGNIFICANT CORRELATIONS BETWEEN NUMBERS OF SIBLINGS OF EACH SEX AND RESPONSES TO DEPENDENT MEASURES FOR MALES, FEMALES AND BOTH COMBINED (ALL Ps < 0·05)

Sex of Subject	Male			Female			Male and Female		
Number of Siblings *	B	S	B+S	B	S	B+S	B	S	B+S
Items									
Own IQ		−0·29							
Own Sociability							−0·20		
Girls work harder		0·28							
Boys work harder		−0·32		0·30	−0·29		0·20	−0·23	
Both work equally				−0·30					
Ideal occupation					0·34			0·19	
O-level grades								0·21	
Arts O-levels			−0·31					−0·20	−0·19
Arts grades (O)				−0·33				0·19	
Science O-levels				−0·38					

* B stands for number of brothers, S for number of sisters, and B + S for total number of brothers and sisters combined.

increased subjects of both sexes were less likely to regard themselves as being sociable and more likely to believe that boys work harder than girls. However, when looking at male and female subjects separately, it was found that while there were no significant correlations between dependent measures and number of brothers for males there were several among females. Females with more brothers were more inclined to believe that boys work harder than girls, less inclined to regard the two sexes as working equally hard, had taken fewer science O-levels, and had performed more poorly at arts O-levels.

There were also several significant overall effects of number of sisters. Across both sexes, those with more sisters were less inclined to believe that boys work harder than girls, this relationship being slightly stronger for male subjects than for females. Those with more sisters also aspired to lower ideal occupational status, an effect which only remained clearly significant among females. Number of sisters was negatively correlated with number of arts O-levels taken but positively with O-level grades. Among males, but not females, those with more sisters were more likely to believe that girls work harder than boys and less likely to regard themselves as having above average IQ.

DISCUSSION

The data indicate sex differences in the expected direction for many of the variables observed. Males study more science and fewer arts subjects, are strikingly more self-confident, and have a more positive academic self-concept than do females. Since *actual* performance (at O-levels) does not differ between the sexes in either science or arts subjects, it appears that the academic self-concepts of these pupils are not simply determined by academic achievement.

Within a social identification framework these results suggest that females may be pushed into a non-academic self-concept owing to the nature of the *intergroup* relations with males. Females seek dimensions of intergroup comparison which do not place them at a disadvantage and, therefore, tend to withdraw from making academic comparisons with males.

If such an interpretation is correct, it follows that females whose gender identity is more salient will be more likely to abandon academic aspirations. As predicted, females were most clearly affected by the number of male siblings in their family, but not by *total* number of siblings. Females' academic aspirations are therefore particularly susceptible to influence by male siblings. It is also interesting that the belief that boys are harder working diminishes as more sisters enter the family, emphasising even more that variation in such beliefs may be the products of intergroup comparisons.

For male subjects increasing numbers of brothers have no impact. In line with our hypotheses it is the number of sisters which seems to be of greatest importance. Here it is not academic *behaviour* which is affected, but relative self-evaluation. Those with more sisters are less likely to think of themselves as above average intelligence, or of boys as working harder than girls. This suggests that having more sisters not only raises the salience of gender among males, but also provides evidence which reduces the stereotypicality of their perceptions of the academic capabilities of the two sexes.

A more complex interpretation than has hitherto been offered also presents itself. For both sexes, increasing numbers of the opposite sex siblings inhibits endorsement of the view that the opposite sex work less hard (are lazier) than own. This suggests that, while making gender salient, minority status among siblings has a different effect from minority status *per se*. Here, Stephenson's (1981, 1984) model of orthogonal intergroup and interpersonal dimensions of behaviour may be of use. Finding oneself to be the only member of one's sex in a gathering of people will tend to make one's gender salient and may lead to intergroup comparisons with the others present. However, if those others are one's siblings *interpersonal* relations and comparisons will cross-cut the intergroup dimension (see discussion of 'imitation' and 'contrast' effects in Wilkinson *et al.* (1982)). Within the family, norms arise to govern and ensure the smooth running of interpersonal relationships between sex groups. These norms, on the whole, ameliorate the more extreme expressions of intergroup attitudes (cf. Stroebe and Stroebe's (1984) treatment of the marital couple as a social group). For the dominant and secure group (males) there is leeway to concede some ground in intergroup attitudes without compromising overall superior ingroup status, or compromising oneself behaviourally (e.g., in the choice of subjects studied). For the insecure subordinate group (females) capitulation to the already dominant stereotype of males as being harder working seems to coincide with a behavioural withdrawal from competing in certain parts of the academic arena. It seems plausible that this is a strategy which allows them to gain positive distinctiveness on alternative dimensions of comparison.

Although the sample and number of subjects are restricted, the data provide a degree of support for the distinctiveness principle as an explanation of identity salience (McGuire and McGuire, 1982). There is also some support for the idea that the *effects* of gender salience may include an important intergroup dimension (Abrams, 1983; Abrams and Condor, 1984; Condor and Abrams, 1984; Huici, 1984) — a consideration which does not figure prominently in the McGuires' approach.

The present findings render less tenable the view that sex-of-sibling effects may be restricted to early childhood (Sutton-Smith and Rosenberg, 1970), and suggest that important effects may arise in adolescence (cf. Kidwell, 1982). Even when direct influence (by instruction, interpersonal conflict, and modelling) has diminished, the sex of one's siblings continues to have impact on salience of one's own gender, which influences the likelihood of intergroup comparisons between own and other sex. These processes are particularly likely to occur during adolescence (Abrams and Condor, 1984; Sherif, 1984) since this is a period of active construction of identity (Coleman, 1974, 1980; Erikson, 1968). The present data serve to reinforce our view that sex of siblings and intergroup relations between the sexes should be the subject of further study. In many cases some of the data are available — for example the CHES study (Osborn *et al.*, 1984) has collected sex of sibling information — but have yet to be analysed within this framework. In others the impact of these factors has simply been ignored.

ACKNOWLEDGMENTS. — We wish to thank Howard Giles and Karen Henwood for their helpful comments on an earlier version of this article. The research reported here was conducted by the second author as part of her undergraduate dissertation, while all three authors were at the Department of Psychology, University of Bristol. Requests for reprints should be addressed to Dominic Abrams, Department of Psychology, The University, Dundee DD1 4HN, Scotland.

REFERENCES

ABRAMS, D. (1983) Light dawns on the children. *New Society*, 65, 216.

ABRAMS, D., and CONDOR, S. (1984). A social identity approach to the development of sex-identification in adolescence. Paper presented at the International Conference on Self and Identity, Cardiff, July.

ALPER, T. G. (1973). The relationship between role-orientation and achievement motivation in college women. *J. Personality*, 41, 9-13

ANYON, J. (1983). Intersections of gender and class: accommodation and resistance by working class and affluent females to contradictory sex-role ideologies. In WALKER, S., and BARTON, L. (Eds.), *Gender, Class and Education*. Lewes: Falmer Press.

BELMONT, L., and MAROLLA, F. A (1973). Birth order, family size and intelligence. *Science*, 182, 1096-1101.

Blackstone, T. (1976) The education of girls today In Mitchell, J., and Oakley, A. (Eds.), *The Rights and Wrongs of Women*. Harmondsworth: Penguin.

Coleman, J. C. (1974). *Relationships in Adolescence*. London: Routledge and Kegan Paul.

Coleman, J. C. (1980). *The Nature of Adolescence*. London: Methuen.

Condor, S., and Abrams, D. (1984). Womanhood as an aspect of social identity: group identification and ideology. Paper presented at the International Conference of Self and Identity, Cardiff, July.

Del Boca, F. L., and Ashmore, R. D. (1981). Sex stereotypes through the life cycle. In Wheeler, L. (Ed.), *Review of Personality and Social Psychology, Vol. 1*. London: Sage.

Dunn, J. (1984). Early social interaction and the development of emotional understanding. In Tajfel, H. (Ed.), *The Social Dimension, Vol. 1*. London: Cambridge University Press.

Dunn, J., and Kendrick, C. (1982). *Siblings: Love, Envy and Understanding*. London: Grant-McIntyre.

Educational Statistics for the United Kingdom, 1983. London: HMSO

Educational Statistics for the United Kingdom, 1984. London: HMSO.

Erikson, E. (1968). *Identity, Youth and Crisis*. London: Faber and Faber.

Horner, M. S. (1972). Towards an understanding of achievement-related conflicts in women. *J. social Issues*, 28, 157-176.

Huici, C. (1984). The individual and social functions of sex role stereotypes. In Tajfel, H (Ed.), *The Social Dimension, Vol. 2*. London: Cambridge University Press.

Katz, P. (1979). The development of female identity. *Sex Roles*, 5, 155-178.

Kidwell, J. S. (1982). The neglected birth order: middleborns. *J. Marriage and the Family*, 44, 225-235.

King, R. (1971). Unequal access in education — sex and social class. *Social and Economic Administration*, 5, 167-275.

Koch, H. L. (1955). Some personality correlates of sex, sibling position and sex of siblings among 5- and 6-year-old children. *Genet. Psychol. Monog.*, 52, 3-50.

Leventhal, G. S. (1970). Influence of brothers and sisters on sex-role behaviour. *J. Person. soc Psychol.*, 16, 452-465.

Marjoribanks, K. (1976) Sibsize, family environment, cognitive performance and effective characteristics. *J. Psychol.*, 94, 195-204.

McGuire, W. J , and McGuire, C. V. (1981). The spontaneous self-concept as affected by personal distinctiveness. In Lynch, M D., Norem-Heberson, A. A., and Gergen, K. J. (Eds.), *Self-Concept. Advances in Theory and Research*, Cambridge, Mass.: Balinger.

McGuire, W. J., and McGuire, C. V. (1982). Significant others in the self-space: sex differences and developmental trends in the social self In Suls, J. (Ed.), *Psychological Perspectives on the Self*. Hillsdale, N.J.: Erlbaum.

McGuire, W. J., and McGuire, C. V. (1984). Developmental trends and gender differences in contents and processes within the phenomenal self. Paper presented at the International Conference on Self and Identity, Cardiff, July.

McGuire, W. J., McGuire, C. V., and Winton, W. (1979). Effects of household sex composition on the salience of one's gender in the spontaneous self-concept. *J. exp. soc. Psychol.*, 15, 77-90.

Osborn, A., Butler, N., and Morris, A. C. (1984). *The Social Life of Britain's 5-year-olds*. London: Routledge and Kegan Paul.

Purvis, J. (1981). The double burden of class and gender in the schooling of working class girls in nineteenth century England 1800–1870. In Barton, L., and Walker, S. (Eds), *Schools, Teachers and Teaching*. Lewes: Falmer Press.

Rosenberg, B. G., and Sutton-Smith, B. (1968). Family interaction effects on masculinity-femininity *J. Person. soc. Psychol.*, 8, 117-121.

Sarah, E., Scott, M., and Spender, D. (1980). The education of feminists: the case for single sex schools. In Spender, D., and Sarah, E. (Eds), *Learning to Lose: Sexism and Education*. London: The Women's Press.

Sharpe, S. (1976). *Just Like a Girl*. Harmondsworth: Penguin.

Sherif, C. W. (1982). Needed concepts in the study of gender identity. *Psychol Women Q.*, 6, 375-398.

Sherif, C. W. (1984). Co-ordinating the sociological and psychological in adolescent interactions. In Doise, W., and Palmonari, A. (Eds.), *Social Interaction in Individual Development*. London: Cambridge University Press.

Stephenson, G. M. (1981). Intergroup bargaining and negotiation. In Turner, J. C., and Giles, H. (Eds.), *Intergroup Behaviour*. Oxford: Blackwell.

Stephenson, G. M. (1984). Intergroup and interpersonal dimensions of bargaining and negotiation. In Tajfel, H. (Ed.), *The Social Dimension, Vol 2*. London: Cambridge University Press.

Stroebe, W., and Stroebe, M. S. (1984). When love dies: an integration of attraction and bereavement research. In Tajfel, H. (Ed.), *The Social Dimension, Vol. 1*. London: Cambridge University Press.

Sutherland, M. B. (1981). *Sex Bias in Education*. Oxford: Blackwell

Sutton-Smith, D , and Rosenberg, B. G. (1970). *The Sibling*. New York: Holt.

Tajfel, H (1984). Intergroup relations, social myths and social justice in social psychology. In Tajfel, H. (Ed.), *The Social Dimension, Vol. 2*. London: Cambridge University Press.

Tajfel, H., and Turner, J. C. (1979). An integrative theory of intergroup conflict. In Austin, W. G., and Worchel, S. (Eds.), *The Social Psychology of Intergroup Relations*. Monterey, Calif.: Brooks-Cole.

TAYLOR, S. E. (1981). A categorisation approach to stereotyping In HAMILTON, D. L. (Ed.), *Cognitive Processes in Stereotyping and Intergroup Behaviour*. Hillsdale, N.J.: Erlbaum.

TAYLOR, S. E., FISKE, S. T., ETCOFF, N. L., and RUDERMAN, A. J. (1978). Categorical and contextual bases of person memory and stereotyping. *J. Person. soc. Psychol.*, 36, 778-793.

TERHUNE, K. W. (1974). A review of actual and expected consequences of family size. U.S. Public Health Service Publications, NIH 75-779; Washington DC., U.S. Government Printing Office.

TURNER, J. C (1981). The experimental social psychology of intergroup behaviour. In TURNER, J. C , and GILES, H. (Eds.), *Intergroup Behaviour*. Oxford: Blackwell.

TURNER, J. C (1982). Towards a cognitive re-definition of the social group. In TAJFEL, H. (Ed.), *Social Identity and Intergroup Relations*. London: Cambridge University Press.

WAGNER, M. A., SCHUBERT, H. J. P., and SCHUBERT, D. S. P (1979). Sibling-constellation effects on psychosocial development, creativity and health. *Advances in Child Development and Behaviour*, 14, 57-148.

WALKER, S., and BARTON, L (Eds.) (1983). *Gender, Class and Education*. Lewes: Falmer Press.

WEINREICH-HASTE, H. (1984). Cynical boys, determined girls? Success and failure anxiety in British adolescents *Br J. soc. Psychol.*, 23, 257-264.

WILKINSON, K., STITT, B. G., and ERICKSON, M. L. (1982). Siblings and delinquent behaviour: an exploratory study of a neglected family variable. *Criminology*, 20, 223-239.

WILLIAMS, J. A., and GILES, H. (1978). The changing status of women in society: an intergroup differentiation perspective. In TAJFEL, H (Ed), *Differentiation Between Social Groups*. London. Academic Press.

ZAJONC, R. B., and MARKUS, G. B (1975). Birth order and intellectual development. *Psychol. Rev.*, 82, 74-88.

(*Manuscript received* 14*th January*, 1985)

Br. J. educ. Psychol., 55, 233-239, 1985

THE RELATIONSHIP OF GRADE, SEX, ABILITY AND SOCIO-ECONOMIC STATUS TO PARENT, PEER AND SCHOOL AFFILIATION

By P. G. RICHMOND
(*School of Education, Macquarie University, N.S.W.*)

SUMMARY. A measure of affiliation to parents, peers and school was administered to a sample of 1379 subjects of both sexes, aged between 11 and 17 years. A principal component analysis of the scores on the inventory produced three clearly separated components, one for each reference group. The scores derived from a rotated solution were analysed using grade, sex, ability, and socio-economic status as independent variables. The main findings were: (i) Home and school affiliation scores declined with age while peer affiliation scores increased, (ii) Female students were more positive about peer and school affiliation than were male students, while male students were more positive about home affiliation, (iii) The high ability students were more positive about their schools than were the low ability students, and (iv) Socio-economic level was not related to reference group affiliation.

INTRODUCTION

EVIDENCE from various studies (Utech and Hoving, 1969; Hamm and Hoving, 1971; Curtis, 1975; Bixenstine *et al.*, 1976; Hunter and Youniss, 1982) indicates that, in the years of late childhood and adolescence, there is a trend towards closer peer affiliation while family and parental influences tend to decrease. The tenor of some results suggests that these two trends are not conflicting (Bowerman and Kinch, 1959; Brittain, 1963), although Berndt (1979) suggests that the influence of peers and parents may be opposed. Other evidence (McConnell, 1963; Costanzo and Shaw, 1966; Landsbaum and Willis, 1971) indicates that alongside the shift of emphasis from parents to peers there is a general decline of influence from both groups as the adolescent becomes more confident and autonomous.

Changes in attitude to school also occur in this period. Some adolescents develop strong feelings of alienation (Hargreaves, 1967, 1972), and in general older adolescents express higher levels of dissatisfaction than do younger adolescents (Musgrove, 1966). Some sex differences in levels of affiliation have been found. Curtis (1975) has shown that females tend to be more positive towards peers than are males, and Montemayor (1982) suggests that for females, but not males, peer and parent influences may be in competition. Skaalvik (1983) indicates that grade 8 females show a significant correlation between academic achievement and perceived value of the school while this relationship is not present for males.

Evidence concerning socio-economic status and reference group affiliation is limited. Curtis (1975) suggests that higher socio-economic groups tend to be more parent oriented. In regard to ability level the work of Hargreaves (1967, 1972) suggests indirectly that a positive view of schooling is more likely in higher ability groups rather than in lower ability groups. The work of Lacey (1970) concerning the importance of student attitudes when selection for higher or lower streams is determined points to the same conclusion, although the study by Skaalvik (1983) suggests that sex differences may be present.

It is the aim of this study to confirm these findings, and to extend them in several ways with respect to the variables of grade, sex, ability and socio-economic status (SES). It is hypothesised on two related grounds that females will have a more positive attitude to school than males. In the first place, following Sexton (1969), it can be argued that the feminine sex role is more compatible with the school ethos than the male sex role. On this basis, female students will be more adjusted and comfortable in the school setting than males. In the second place, the work of Kulka *et al.* (1982) and Lagerspetz *et al.* (1982) shows that aggressiveness in males is associated with negative attitudes towards school staff and school rules. On the ground that males tend to be more aggressive than females they might be expected to show a less positive attitude to school.

It can be hypothesised, also, that ability will be an important variable in determining a student's attitude to school, such that high achieving students will feel more positive about school than will low achieving students. In addition, high achieving students should be less peer oriented since high achievement is often attained in competition with and at the expense of peers. On the other hand, low achievers may turn to their peers for support when school appears to reject them. In regard to SES, it is hypothesised that students from higher SES backgrounds will show raised levels of school affiliation on the grounds that in general there is greater compatibility of attitudes between school and high SES students than between school and low SES students.

The study to be described examines the levels of affiliation to parents, to peers, and to school as expressed by students aged between 11 years and 17 years, taking into account the grade level, sex, ability and socio-economic status of the sample.

METHOD

Measure of group affiliation

An inventory of 15 items was selected to measure group affiliation. Two pilot studies, containing about 300 subjects in each, were used to develop the inventory. The psychometric properties of the items were determined and modifications were made to improve the wording and to increase reliability and consistency. The measure was allied to a 6-point scale of the following form: Never or almost never (1), Once in a while (2), Some of the time (3), Often (4), Most of the time (5), Always or almost always (6). The inventory also contained five items which provide a means for assessing the acceptability of the responses of each subject. The items were: I do my best; I behave well; I do the right things; I tell the truth; I do as I am told. These items provide a means of assessing response distortion (Goodstein and Lanyon, 1971, chapter 7). The items may be used in two ways. First to identify respondents who present themselves in an unbelievably poor light, i.e., Heilbrun's (1964) concept of "faking bad", and second, to identify respondents who present themselves in an unbelievably good light, i.e., Heilbrun's concept of "faking good".

The subjects

The sample contained 1468 subjects with approximately equal numbers of each sex. They were members of the sixth, eighth and tenth grade classes in eight co-educational schools in metropolitan Sydney, Australia. The number of subjects in each grade was 374, 569 and 525 respectively, and the approximate average age of the grade groups was 12, 14 and 16 years. Congalton (1969) was used to select four schools from high status suburbs and four schools from low status suburbs. The Congalton scale (1969, pp. 5-7) was developed from the responses of a sample of Sydney residents who were asked to rank the various suburban areas on a seven-point scale ranging from (1) highest social standing, to (7) lowest social standing. When Congalton's index was applied to the sample schools a value of 2·5 was obtained for the upper status suburbs and a value of 5·3 was obtained for the lower status suburbs. The differences between the suburbs can be defined in terms of such criteria as income levels, educational levels of parents, unemployment levels, occupational status and quality of housing (Poulsen and Spearritt, 1981). It is possible that the division of the sample in this way will produce bias at the 10th grade level. Differences in drop-out rate between various local government regions arising from early school leaving have been demonstrated by Connell *et al.* (1975). The drop-out rate tends to be greater in low social status areas. In consequence, the tenth grade sample in the lower social status areas might be untypically favourable to schooling because the students with unfavourable attitudes have left. To check upon this, the number of early leaving students in the 10th grade of each school was assessed. It was found that in the lower social status areas a 4 per cent drop-out occurred, while the rate was less that 1 per cent in the higher status areas. It was concluded that the drop-out rates at the time of sampling were not marked enough to produce a significant bias.

Two ability groups were formed in the following way. In the secondary schools the criterion for high ability was membership of the top streamed classes in mathematics. Low ability was defined by membership of the low streamed classes in mathematics. In the primary schools, the class teachers were asked to categorise their children as high or low in attainment using the basic subjects as the criterion.

PROCEDURE AND RESULTS

The inventory was administered to class-sized groups and each item was scored using the 6-point scale. Case acceptability was monitored using the five items listed above. Subjects were removed from the sample if their total score on these items was outside the range ± 2 SD for the sample. The final sample contained 1379 subjects of whom 683 were female and 696 were male. Cronbach's alphas for the sub-measures of home affiliation, school affiliation and peer affiliation were 0·82, 0·74, and 0·65 respectively.

Component analysis

A principal component analysis of the 15 affiliation items was performed with a varimax rotation for the first three components. These components had eigen values greater than unity, and the percentage of variance extracted fell away after the third

TABLE I

UNROTATED COMPONENT SOLUTION FOR THE FIFTEEN AFFILIATION ITEMS

	Component			
Item	1	2	3	h^2
My parents understand me	0·674	−0·362	−0·338	0·700
My parents consider my feelings	0·624	−0·330	−0·392	0·653
My parents and I have fun together	0·682	−0·261	−0·228	0·586
I get plenty of attention at home	0·559	−0·148	−0·391	0·488
My friends pick on me	−0·134	−0·386	0·246	0·227
I get along well with my school work	0·644	0·053	0·498	0·667
I do the best work I can at school	0·558	−0·005	0·502	0·563
I am proud of my school work	0·594	0·070	0·466	0·575
I like being at school	0·535	0·014	0·390	0·439
I feel upset at school	−0·419	−0·008	−0·129	0·193
I am popular with my friends	0·349	0·645	−0·214	0·584
I am a lot of fun to be with	0·365	0·620	−0·113	0·530
I am easy to like	0·388	0·557	−0·157	0·485
I like being part of the group	0·258	0·430	−0·160	0·277
I would like to get away from home	−0·579	0·352	0·074	0·465
Eigenvalue	3·99	1·90	1·53	—

TABLE 2

VARIMAX ROTATION FOR THE FIFTEEN AFFILIATION ITEMS DERIVED FROM THE SOLUTION SHOWN IN TABLE 1

Items	Rotated Components		
	1	2	3
My parents understand me	0·821	0·157	0·030
My parents consider my feelings	0·801	0·084	0·059
My parents and I have fun together	0·718	0·251	0·083
I get plenty of attention at home	0·668	0·052	0·195
I would like to get away from home	−0·607	−0·298	0·083
I get along well with my school work	0·144	0·797	0·097
I do the best work I can at school	0·110	0·742	0·015
I am proud of my school work	0·119	0·741	0·106
I like being at school	0·147	0·643	0·062
I feel upset at school	−0·212	−0·369	−0·107
I am popular with my friends	0·045	0·090	0·757
I am a lot of fun to be with	0·013	0·177	0·706
I am easy to like	0·083	0·155	0·674
I like being part of the group	0·056	0·063	0·519
My friends pick on me	−0·039	0·083	−0·468
% Variance	18·5	17·0	14·1

component. The variance contained in the first three components was 49·6 per cent. The unrotated matrix for the first three components is shown in Table 1, and the varimax rotated solution is shown in Table 2.

It can be seen in Table 1 that the first component is general and positive, with the exception of three loadings which derive from items which have reversed scaling. The second component is bipolar. The positive pole is defined by four peer items together with the one reversed parent item, while at the negative pole there are four parent items together with one reversed peer item. The school items have negligible loadings on this component. In the third component, the school items together with the reversed peer item are contrasted with the remaining items. The rotated varimax solution is shown in Table 2. It can be seen that the first component contains large loadings from the parent items. The remaining loadings are negligible. The second component contains large loadings from the school items, and small to negligible loadings from the other items. The third component contains large loadings from the peer items, and negligible loadings from the remainder. It appears that the rotated solution has three well defined components with little or no spillage of loadings between the components. Component scores for each subject were computed from both the unrotated solution in Table 1 and the rotated solution shown in Table 2. These scores were used in the analysis of variance discussed in the next section.

Analysis of variance

An analysis of variance was performed upon each set of rotated component scores shown in Table 2. Grade, sex, ability and socio-economic status (SES) were used as independent variables. The results obtained are shown in Table 3.

Table 3 indicates that there were differences in parent affiliation scores according to grade. For both sexes, affiliation scores declined across grades 6, 8 and 10. A sex difference was also present such that males showed higher scores for affiliation with parents than did females. This pattern was consistent at each grade, although it was most marked at the 8th grade. Differences in school affiliation scores occurred according to ability. For both sexes and across all grades, the higher ability groups

TABLE 3

SUMMARY OF SIGNIFICANT RESULTS OBTAINED IN THE ANALYSIS OF VARIANCE, USING THE ROTATED COMPONENT SCORES

Dependent Variable and Source of Variation[1]	s.s.	df	F Value	Sig of Value
Parent Affiliation				
Grade	69·7051	2	36·77	0 000
Sex	6·1022	1	6·44	0 01
School Affiliation				
Ability	59·43454	1	65·03	0 000
Grade	43 0283	2	23·54	0·000
Sex	12 5230	1	13·70	0 000
Grade × SES[2]	9 2494	1	5·06	0 006
Peer Affiliation				
Grade	22 9993	2	11·96	0·000
Sex	11·8601	1	12·33	0·000
Grade × SES[2]	17·1838	2	8 93	0·000
Grade × Ability	8·2895	2	4·31	0·01

[1] Only significant effects and interactions are shown although all the independent variables and their interactions were included in each analysis

[2] SES = Socio-economic status.

showed larger school affiliation scores than did the lower ability groups. For both sexes school affiliation scores declined across grades 6, 8 and 10 regardless of ability level. In addition, at each grade, females showed higher school affiliation scores than did males. An interaction between grade and socio-economic status was significant for school affiliation, such that at 6th grade, the high SES students produced higher school affiliation scores than the low SES students, while at 10th grade, this pattern was reversed. Peer affiliation scores increased across grades 6, 8 and 10 for both sexes. At 6th grade, there was no sex difference, while at 8th the females showed higher affiliation scores and at 10th grade the difference was more marked. An interaction between grade and SES was significant for peer affiliation. For 8th grade students, the high SES group showed raised peer affiliation scores, while at 6th and 10th grade the low SES group showed raised scores. A significant interaction also occurred between grade and ability. In 6th grade, the more able students showed raised peer affiliation scores, while at 8th and 10th grade, the less able showed raised peer affiliation scores.

A limited analysis of the unrotated component scores was carried out. The first unrotated component shown in Table 1 was treated as a dependent variable with grade as an independent variable. Positive attitudes towards the three contexts declined across the grades at $P<0{\cdot}001$ (s.s. 68·6452; df 2; $F=36{\cdot}07$). The second unrotated component was examined for sex differences and it was found that female scores tended towards the positive pole (peer scores greater than parent scores) while male scores tended towards the negative pole (parent scores greater than peer scores) at $P<0{\cdot}001$ (s.s. 23·4275; df 1; $F=23{\cdot}81$).

DISCUSSION

Shifts in affiliation between parents, peers and school

The results indicate that between the ages of 12 and 16 years, for both sexes, there is a decline in affinity with parents and an increase in affinity with peers. The finding is in line with a substantial body of previous work. Researchers are not agreed, however, about whether the cross pressures between parents and peers constitute conflict for the adolescent. The unrotated component solution in Table 1 provides evidence relevant to this issue. The first unrotated component is general, which indicates that members of the sample held attitudes which were common to all three reference groups. The second component is bipolar, and following Burt (1940), this indicates the presence in the data of two "contrasting tendencies", that of parent affiliation and of peer affiliation. At the positive pole, the scores derived from this component will be those of the individuals with large difference scores favouring peer affiliation, while at the negative pole the difference scores will favour parent affiliation. The study indicates that the most marked natural contrast in the data, after the extraction of a general affiliative attitude, derives from the difference between parent and peer affiliation scores. Thus, there is evidence of a contrast between peer and parent affinity. The finding that school affiliation declines across grades 6, 8 and 10 supports the work of Musgrove (1966). Overall, the general affiliative attitude which is represented in the first unrotated component shown in Table 1, shows a significant decline between grade 6 and grade 10.

Sex differences

Several sex differences were found. Secondary school male students were more parent and less peer oriented than were their female counterparts. In terms of parent–peer contrast scores, the females showed differences tending to favour peers while the males showed differences tending to favour parents. This result supports the work of Curtis (1975), and since the parent–peer contrast score for females tends to favour peers, the result can be taken to support the work of Montemayor (1982). At all grades, the females were more positive towards their school than were the males. This result is in the predicted direction and can be explained on the grounds that feminine attitudes are more in tune with the school ethos.

Ability differences

The able students showed a more positive attitude to school than did the less able students, regardless of sex or grade level. This result was in the predicted direction. Skaalvik's (1983) finding concerning a sex difference at 8th grade was not confirmed. There were no differences in parent or peer affiliation which were related to ability, although for peer affiliation a significant interaction between ability and grade was found. At 8th and 10th grades, the more able students had lower peer affiliation scores than did the less able students. This result was in the predicted direction. However, the reverse pattern occurred at grade 6, such that high ability students showed higher levels of peer affiliation. This outcome may reflect differences in school ethos and organisation between the primary and secondary levels of schooling.

Socio-economic differences

No significant differences occurred with any of the affiliation measures in relation to the socio-economic variable. The prediction that high SES students would be more positive towards school was not confirmed. In fact, a significant interaction occurred in which the opposite relationship was present in the 10th grade. It is possible that the small socio-economic differences in the rates of early leaving which were found in the 10th grade year were, after all, an influential factor in this outcome. In regard to peer affiliation, grade 8 students showed an opposite pattern to the rest of the sample. At this grade, the high SES group showed higher levels of peer affiliation. A tenable explanation for this interaction is not apparent.

CONCLUSION

By and large, the present findings confirm and bring together a wide range of previous research. Three interesting extensions of the area are worth noting. Firstly, there appears to be a difference between the sexes concerned with their attitude to schooling. Across the age range 12 to 16 years, the female students expressed a more favourable attitude about their schooling than did the male students. It may be the case that sex role characteristics like those of independence, individuality and assertiveness (if not aggression) in males, and compliance, co-operativeness and supportiveness in females, set up differing social interactions between teachers and students. The present findings reinforce the view of Skaalvik (1983) that social interaction processes in school may differ between the sexes, and these differences need to be taken into account when research is planned. Secondly, there is the influence of academic achievement or ability level upon the students' evaluation of school. Students viewed their schools favourably when they were academically successful. Thirdly, there is the finding that in the secondary school, able students showed lower levels of peer affiliation than did less able students, while in the primary school the reverse pattern occurred. This outcome merits further study to establish whether the result can be confirmed, and whether particular aspects of school organisation or climate are responsible.

NOTE. This research was supported by a Macquarie University Research Grant. A copy of the inventory used may be obtained from the author at the School of Education, Macquarie University, North Ryde, NSW 2113, Australia.

REFERENCES

Berndt, T. J. (1979). Developmental changes in conformity to peers and parents. *Dev. Psychol.*, 15, 608-616.

Bixenstine, V. E., DeCorte, M. S., and Bixenstine, B. A. (1976). Conformity to peer sponsored misconduct at four grade levels. *Dev. Psychol.*, 12, 226-236.

Bowerman, C. E., and Kinch, J. W. (1959). Changes in family and peer orientation of children between fourth and tenth grades. *Social Forces*, 37, 206-211.

Brittain, C. V. (1963). Adolescent choices and parent-peer cross-pressures. *Am. sociol. Rev.*, 28, 385-391.

Burt, C. (1940) *The Factors of the Mind*. London: University of London Press.

Congalton, A. A. (1969). *Status and Prestige in Australia*. Melbourne: Cheshire.

Connell, W. F, Stroobant, R E., Sinclair, K. E., Connell, R. W., and Rogers, K W (1975). *12 to 20, Studies of City Youth*. Sydney: Hicks Smith and Sons.

Costanzo, P. R., and Shaw, M. E. (1966). Conformity as a function of age level. *Child Dev.*, 37, 967-975.

Curtis, R. L. (1975). Adolescent orientations towards parents and peers: variations by sex, age and socio-economic status. *Adolescence*, 10, 483-494

Goodstein, L. D., and Lanyon, R. I. (1971). *Readings in Personality Assessment*. New York: J. Wiley.

Hamm, N. H. and Hoving, K. L. (1971). Conformity in children as a function of grade level, and real *versus* hypothetical adult and peer models. *J genet. Psychol.*, 118, 253-263.

Hargreaves, D. H (1967). *Social Relations in a Secondary School*. London· Routledge and Kegan Paul.

Hargreaves, D. H. (1972). *Interpersonal Relations and Education*. London· Routledge and Kegan Paul.

Heilbrun, A B (1964) Social learning theory, social desirability and the MMPI. *Psychol. Bull.*, 61, 377-387.

Hunter, F. T., and Youniss, J (1982). Changes in functions of three relations during adolescence. *Dev. Psychol.*, 18, 806-811

Kulka, R. A., Kahle, L. R., and Klingel, D. M. (1982). Aggression, deviance and personality adaptation as antecedents and consequences of alienation and involvement in high school. *J. Youth Adolesc.*. 11, 261-279.

Lacey, C. (1970). *Hightown Grammar*. Manchester: Manchester University Press

Lagerspetz, K M. J., Bjorkqvist, K, Berts, M., and King, E. (1982). Group aggression among school children, in three schools *Scand. J. Psychol.*, 23, 45-52.

Landsbaum, J B., and Willis, R. H. (1971). Conformity in early and late adolescence. *Dev. Psychol.*, 4, 334-337.

McConnell, T. R. (1963). Suggestibility in children as a function of chronological age. *J. abnorm soc. Psychol.*, 67, 286-289.

Montemayor, R. (1982). The relationship between parent-adolescent conflict and the amount of time adolescents spend alone and with parents and peers *Child Dev.*, 53, 1512-1519.

Musgrove, F. (1966). The social needs and satisfactions of some young people. Part II — At school. *Br. J. educ. Psychol*, 36, 137-149.

Poulsen, M., and Spearritt, P. (1981). *Sydney, a Social and Political Atlas*. Sydney: George Allen and Unwin.

Sexton, P. C. (1969). *The Feminized Male*. New York. Random House.

Skaalvik, E. M (1983). Academic achievement, self esteem and valuing of the school — some sex differences. *Br. J. educ. Psychol.*, 53, 299-306.

Utech, D A, and Hoving, K L. (1969). Parents and peers as competing influences in the decisions of children of differing ages *J. soc Psychol.*, 78, 267-274.

(*Manuscript received* 4*th March*, 1985)

Br. J. educ. Psychol., 55, 240-250, 1985

FURTHER EVIDENCE FOR GENETIC INFLUENCES ON EDUCATIONAL ACHIEVEMENT

By C. E. GILL
(*Biometrics Unit, University of Western Australia*)
R. JARDINE
(*Research School of Biological Sciences, Australian National University*)
AND N. G. MARTIN
(*Department of Human Genetics, Medical College of Virginia, USA*)

SUMMARY. A genetic analysis of Tertiary Admission Examination (TAE) and Australian Scholastic Aptitude Test (ASAT) results for 264 pairs of MZ and DZ twins is reported. Purely environmental models are rejected as inadequate explanations of variation in examination performance and genetic factors must be invoked to obtain a satisfactory fit to the data. Within the portion of the age cohort who are candidates for these examinations, genetic factors appear to account for about 70 per cent of variation while environmental experiences shared by siblings appear to have little or no influence. However, when corrections are made on the assumption that examination candidates represent the top 34 per cent from a normal distribution of ability in the population, much greater variation between families is inferred for the population. If we also take account of the high correlation in educational achievement between husbands and wives the putative population twin correlations are consistent with heritabilities between 0.6 and 0.7 and modest contributions of shared environment around 20 per cent of the total variance. The data suggest that a distinction between IQ tests and tests of scholastic achievement on the basis of their causes of variation is not justified. We also show that while a common genetic factor is responsible for much of the covariation of ASAT Arts and Science scores, there are also some differences in the genes responsible for variation in the two areas.

INTRODUCTION

IN the last decade there have been great strides forward in the development of efficient methods for the analysis of individual differences and in their application to cognitive abilities. There is now little doubt that genetic factors contribute significantly to individual differences in intelligence and educational achievement (Eaves, 1975; Rao *et al.*, 1977; Rice *et al.*, 1980; Horn *et al.*, 1982) although it has been suggested that the heritability for scholastic achievement is lower than that for IQ (Jensen, 1969). Of far greater moment than the precise value of the heritability, however, are issues frequently overlooked in the traditional literature on the genetics of cognition. Is there evidence for environmental influences shared by siblings which would support the supposed importance of home background and differences between schools? Are people with different genes for ability equally influenced by environmental factors such as teaching methods (genotypes × environments interaction)? Is all genetic variation additive so that children's abilities can, on average, be predicted from parents' abilities, or is there non-additive genetic variation which makes such a prediction less accurate? What is the effect on ability differences between families of the well-known tendency of people to marry spouses of ability similar to themselves?

The most powerful method for beginning to answer some of these questions is the classical twin study in which monozygotic (MZ) and dizygotic (DZ) twins are compared. Martin (1975), analysing the South Australian Intermediate Examination results of 149 pairs of twins aged 14-15 years, found heritabilities in the range 0.76-0.89 for performance in mathematics, the sciences, languages and geography. He concluded that educational achievements; at least as assessed in large scale testing programmes, have heritabilities in the same range as IQ. One puzzling feature of Martin's results, however, was the absence of any evidence for the between pairs component of variation which one would have expected if cultural transmission and assortative mating were making any contribution to variation in educational achievement, as they do for IQ (Fulker and Eysenck, 1979).

In most studies of educational achievement the less able portion of the cohort escapes ascertainment (Martin, 1975; Loehlin and Nichols, 1976). Martin and Wilson (1982) have shown that one-sided sample selection of this kind can cause considerable bias in the estimation of genetic and environmental components of variance which are applicable to the population as opposed to the truncate sample. When they applied a correction to Martin's (1975) results to allow for the fact that only the top 67 per cent of the cohort were examined, estimates of population variance were now consistent with the existence of between families variance attributable to assortative mating or cultural transmission.

In the present study we analyse examination results in a sample of 17-year-old twins in which only the most able 34 per cent of the age cohort are represented. We attempt to take account of this severe truncate selection of the twin sample and also of the well-known correlation between husbands and wives for educational attainment in our analysis of the causes of individual differences in examination performance. We show that once allowance is made for these two complications the breakdown of variance is very similar to that found for IQ.

A further issue is the cause of correlation between ability scores. Two views may be advanced which are related to the classic debate between Spearman and Guilford on the factor structure of IQ. At one extreme is the view that all correlations are caused by a common set of genes which determine general ability and that it is environmental influences which channel this ability towards excellence in one subject or another. At the other extreme are envisaged independent gene effects, each responsible for variation in a different ability while environmental circumstances cause correlation between them by moving all scores up or down in concert. In this paper we attempt to throw light on this issue by analysing the causes of covariation between scores on an Arts and a Science test.

SAMPLE AND MEASUREMENTS

Each year approximately 8700 final year high school students (average age 17 years), or 34 per cent of the age cohort, attempt the Tertiary Admissions Examination (TAE) in Western Australia. There is a choice of 35 subjects in which a candidate may be examined although many of these are attempted by very few people. Candidates attempting the TAE normally enter for six subjects, with a small number attempting seven or more. Since 1976, most candidates have also completed the Australian Scholastic Aptitude Test (ASAT), developed by the Australian Council for Educational Research, which yields scores for performance in Mathematics/Science (ASAT Science) and Humanities/Social Science (ASAT Arts). To ensure comparability between examination results in different subjects, TAE scores are scaled by an anchor variable method, whereby the raw scores of a group taking a particular subject are scaled according to their performance on the ASAT test.

Through the Tertiary Institutions Services Centre, which administers the TAE, the magnetic tapes containing the TAE and ASAT results of all candidates in the period from 1975 to 1981 were used to extract candidates with the same surname and birthdate who had been enrolled in a particular year. This generated 517 pairs of whom 267 pairs shared the same address, 8 pairs repeated the examination and were listed twice, and the remaining 242 pairs were at different addresses. Each of the 509 prospective twin pairs was contracted to see if they were twins, and whether they were willing to cooperate in the study. No contact could be established with five pairs. Of the 242 pairs at different addresses only two pairs were found to be twins. Our final sample for which TAE and ASAT results were available thus consisted of 264 pairs of twins (41 MZ males, 55 MZ females, 52 DZ males, 38 DZ females, 78 DZ opposite-sex). Because many subjects were only attempted by a small number of candidates, we have analysed only those subjects where at least 90 twin pairs were concordant for subject choice, (English, Geography, Mathematics I and Biology), which is probably the lower limit for the useful application of the statistical methods employed below (Martin *et al.*, 1978).

Diagnosis of the zygosity of same-sex pairs was based on their response to the following questions:

1. As children were you and your twin mistaken by people who knew you?

(*a*) Frequently
(*b*) Sometimes
(*c*) Rarely

2. "Non-identical twins are no more alike than ordinary brothers and sisters. Identical twins on the other hand have such a strong resemblance to each other in stature, colouring, features of the face, etc. that people often mistake one for the other" Having read the above statement, do you think you are?

(*a*) Identical
(*b*) Non-identical

This method of zygosity diagnosis has been found by other workers (Cederlöf *et al.*, 1961; Nichols and Bilbro, 1966; Martin and Martin, 1975; Kasriel and Eaves, 1976) to be about 95 per cent correct, approximately the same reliability as obtained by typing for the most common six or seven blood group polymorphisms.

RESULTS

Sampling

Ideally one would like to analyse the data separately for each year but the numbers are too small to make this practicable. It is necessary, therefore, to pool the results of twins over years, first checking that their means and variances are homogeneous. To assess any overall significant differences over years in the means of the TAE and ASAT results for twins (Table 1), separate one-way analyses of variance were performed. Multiple comparisons were made using Scheffe's test. With the exception of Mathematics I, where there is a significant difference between the 1976 and 1977 values, there is no significant difference between the subject means for the different years. Homogeneity of the twin total variances across years (Table 1) was tested using the Bartlett-Box F test and no significant differences were found.

Since the purpose of a genetical analysis of twin data is to make inferences about the causes of variation in the population, it is important that twins are comparable with non-twin subjects. Table 2 presents the means and variances for the total population of candidates attempting the TAE and ASAT in the years 1976 to 1981. Although in the majority of cases the twin sample had a higher mean than that of the total population, the differences were not great. Differences in variances were similarly small. We are thus justified in generalising the results from the twin sample to the population of TAE and ASAT candidates.

Before fitting models to explain trait variation it is important to test whether the individuals in the MZ and DZ groups have been drawn at random from the same population by testing whether the subgroup means and variances are equal. Table 3

TABLE 1

EXAMINATION RESULTS FOR TWIN INDIVIDUALS BY SUBJECT AND YEAR

Subject		Year 1975	1976	1977	1978	1979	1980	1981	All years combined
ASAT Arts	Mean	N/A	26·27	25 80	33·24	28·01	30·91	29·20	28 40
	SD		8·61	7·85	8·65	7·35	7·10	7·80	8 55
	N		93	76	74	69	97	61	470
ASAT Science	Mean	N/A	31·50	28·95	25·84	29·36	36·55	31 67	30·95
	SD		7·07	7·82	8·18	9·65	7·63	9 02	8·81
	N		93	76	74	69	97	61	470
ASAT Total	Mean	N/A	57·76	51·75	59·08	57 38	67·45	60·87	59·35
	SD		14·38	13 82	15·96	15·89	13 80	15·01	11 48
	N		93	76	74	69	97	61	470
English	Mean	60 99	55·77	57·15	56 13	55·30	56·26	60·18	57·03
	SD	16·00	14·93	14·10	13·05	15·71	15·33	13 59	14 73
	N	44	94	73	69	69	91	59	499
Geography	Mean	53·62	53·87	51 40	55·50	51·52	53 44	55·00	53·44
	SD	15·17	15·80	14 50	13·00	15 47	15·75	11·12	14·56
	N	20	58	44	40	35	43	29	269
Mathematics I	Mean	57·24	61·81	51·44	56·35	59·85	57·79	60·32	57·82
	SD	12 35	11·39	12·68	13·25	13 93	15·84	13 20	13 78
	N	22	45	43	39	35	63	39	286
Biology	Mean	51·11	54·58	51·49	54·73	54·08	55 34	55·66	54·09
	SD	12·45	11·63	15·09	10·91	16·35	12·72	12·28	13·16
	N	14	47	46	47	40	45	34	273

lists the means and variances of the TAE and ASAT scores for the five twin groups. Two-tailed tests and variance ratio tests were performed between MZ and DZ means and total variances, separately for males and females. No significant differences in means or total variances in either sex were found.

It is of interest to know whether there is a difference between MZ and DZ twins in their concordance for subject choice. Table 4 shows that MZ twins choose more subjects in common than same-sex DZ twins which suggests that genetic differences influence the choice of subject taken. In his smaller sample of twin pairs Martin (1975) found no significant differences in concordance for subject choice between MZ and DZ same-sex pairs.

Testing hypotheses about the causes of trait variation

Traditional analyses of twin data have done little more than compare the correlations of MZ and DZ twins calculate a "heritability" based on some crude formula which is both inefficient, in that it uses only part of the information available, and often inaccurate in that its calculation is based upon a number of untested assumptions. For these reasons Jinks and Fulker (1970) urged the abandonment of such practices and the rigorous application of the hypothesis testing approach of biometrical genetics. Alternative models of trait variation are fitted to between-and within-pairs mean squares by the method of iterative weighted least squares (Eaves and Eysenck, 1975; Martin, 1975, Eaves *et al.*, 1978). A chi-square test of goodness-of-fit, based on expected mean squares calculated from the least squares parameter estimates, then provides a test of the adequacy of each model.

A large difference in the means of males and females will inflate the within pairs

TABLE 2

EXAMINATION RESULTS BY SUBJECT AND YEAR FOR ALL CANDIDATES

	1977			1978			1979			1980			1981		
Subject	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD
ASAT Arts	8786	23·5	8·5	8612	26·3	8·0	8714	27·7	8·8	8768	35·0	8·7	8678	28·8	10·1
ASAT Science	8786	29·5	7·1	8612	34·7	7·7	8714	27·8	7·0	8768	31·0	6·8	8678	28·1	7·4
ASAT Total	8786	53·0	13·9	8612	61·0	14·3	8714	55·5	14·1	8768	65·9	14·0	8678	56·9	15·5
English	8786	54·9	15·2	8703	57·7	13·8	8803	56·4	15·0	8882	56·7	15·0	8765	56·3	14·9
Geography	3853	52·4	14·5	3810	54·3	13·5	3823	53·3	14·3	3730	52·9	14·6	3525	53·1	14·5
Mathematics I	3849	54·6	14·2	4072	55·6	14·8	4134	55·7	14·5	4128	57·1	14·1	4136	56·8	13·7
Biology	4490	52·3	13·8	4626	53·8	13·3	4579	52·9	14·0	4509	53·6	14·0	3917	53·1	14·0

TABLE 3

SUBJECT MEANS, VARIANCES AND NUMBER OF INDIVIDUALS BY TWIN TYPE, SUMMED OVER YEARS

	MZ Males			MZ Females			DZ Males			DZ Females			DZ Opposite-Sex		
Subject	N	Mean	Variance	N	Mean	Variance	N	Mean	Variance	N	Mean	Variance	N	Mean	Variance
ASAT Arts	66	28·06	72·93	100	29·06	77·44	98	27·84	81·54	66	26·76	59·29	140	29·26	69·89
ASAT Science	66	29·88	80·10	100	31·77	73·27	98	30·41	93·51	66	29·99	67·24	140	31·69	73·10
English	80	52·11	189·61	103	61·37	161·54	95	53·90	238·39	72	59·25	188·79	149	57·58	240·25
Geography	45	50·92	186·05	53	54·71	201·64	52	51·74	243·67	42	53·84	244·30	77	54·96	197·40
Mathematics I	36	55·70	165·64	61	58·87	254·40	50	58·10	243·36	45	55·06	103·84	94	59·14	169·78
Biology	24	51·70	137·59	63	54·96	138·30	43	51·48	199·37	54	54·48	196·00	89	55·15	182·25

TABLE 4

CONCORDANCE OF SUBJECT CHOICE FOR DIFFERENT TYPES OF TWINS

Twin type		Number of subjects in common 0	1	2	3	4	5	6	7	8	N	Mean
MZM	Frequency	0	0	4	7	7	6	14	2	1	41	4·71
MZF		2	2	2	5	9	14	19	2	0	55	4·64
DZM		5	2	4	7	15	11	8	0	0	52	3 73
DZF		0	8	5	6	8	9	2	0	0	38	3·29
DZOS		5	10	9	28	15	7	3	1	0	78	2 97

mean squares (WMS) of DZ opposite-sex pairs. In every case except ASAT Science, female twins have a higher mean score than males and so the DZ opposite-sex within pairs mean squares have been corrected for these sex differences (Clark *et al.*, 1980). The mean squares and their degrees of freedom for each variable are shown in Table 5.

A model for variation in MZ and DZ mean squares is shown in Table 6. E_1 is the environmental variance within families, specific to the individual and shared with no one else, not even members of the same family; it also includes measurement error. E_2,

TABLE 5

OBSERVED MEAN SQUARES BETWEEN (b) AND WITHIN (w) PAIRS USED IN MODEL FITTING

Statistic	df	ASAT Arts	df	ASAT Science	df	English	df	Geography	df	Mathematics I	df	Biology
MZM$_b$	32	129 93	32	139·78	39	324·46	17	290 46	13	302·77	5	165 39
MZM$_w$	33	17·70	33	22·36	40	58 36	18	48 81	14	60 22	6	53 70
MZF$_b$	48	137 00	48	124·57	49	247·64	24	308·01	24	289 77	27	232 50
MZF$_w$	49	17 86	49	18 48	50	65·70	25	92·59	25	204·49	28	58 60
DZM$_b$	48	131 51	48	154 85	45	354 49	19	471 75	16	316·92	13	218·53
DZM$_w$	49	32 55	49	33 49	46	127·33	20	69 59	17	135·26	14	62 34
DZF$_b$	31	84·27	31	94·09	34	228 86	14	250 73	14	183·76	19	329 75
DZF$_w$	32	36·00	32	37 42	35	155·98	15	210·97	15	57 34	20	75·61
DZO$_b$	67	89 88	67	101·29	70	274 58	21	242·18	26	270 73	25	221 34
DZO^a_w	67	48·06	67	46·61	70	177·21	21	156·42	26	95·59	25	127 09

[a] Corrected for mean differences between males and females

TABLE 6

SIMPLE MODEL FOR TWIN MEAN SQUARES

Mean squares		E_1	E_2	V_A
MZ female	Between	1	2	2
	Within	1	0	0
MZ male	Between	1	2	2
	Within	1	0	0
DZ female	Between	1	2	$\frac{3}{2}$
	Within	1	0	$\frac{1}{2}$
DZ male	Between	1	2	$\frac{3}{2}$
	Within	1	0	$\frac{1}{2}$
DZ opposite-sex	Between	1	2	$\frac{3}{2}$
	Within	1	0	$\frac{1}{2}$

the between families environmental component, estimates sources of environmental variance shared by both members of a twin pair but differing between pairs, including such factors as cultural differences and parental rearing practices. Since the great majority of our twin pairs attended the same school E_2 would also include the effect of any differences between school environments.

V_A is that part of genetic variation due to the additive effect of genes in the absence of assortative mating. If assortative mating (the tendency of like to marry like, as measured by the correlation between husbands and wives) occurs for a given character it will increase the additive genetic variance between pairs but in proportions such that it is completely confounded with estimates of E_2 in twin data. Any variation attributed to this source must therefore be interpreted with caution since it could arise from assortative mating, shared environment or both.

In the interests of parsimony the simplest model should be fitted first and more complicated models considered only if simpler ones fail or a significant improvement in fit is obtained by addition of extra parameters. Accordingly we first fit a simple environmental model containing E_1 alone. Failure of this model indicates that there is significant variation between families. A model incorporating E_1 and E_2 tests whether the between-families variation is entirely environmental in origin, whilst an alternative incorporating E_1 and V_A tests whether it is entirely genetic. If both two parameter models fail, a model including E_1, E_2 and V_A must be considered.

Results of fitting models to explain trait variation

The results of fitting models of variation to the TAE and ASAT results are summarised in Table 7. With the exception of Mathematics I, a model (E_1) postulating

TABLE 7

SUMMARY OF RESULTS OF MODEL-FITTING HERITABILITY ESTIMATES GIVEN WHERE APPROPRIATE

Subject	Data Set†	$\hat{E}_1$	$\hat{E}_2$	$\hat{V}_A$	df	χ^2	$\hat{h}^2$
ASAT Arts	10	32 29***	40·67***	—	8	20·91**	
		17·03***	—	54 68***	8	6 01	0 76±0·04
		17·67***	9 04	45·99***	7	5·26	
ASAT Science	10	33 06***	44·50***	—	8	16·26*	
		18 73***	—	57 23***	8	9 31	0 75±0·04
		20·14***	17·81*	39 82***	7	5 99	0 51±0·14
English	10	121·70***	82·04***	—	8	22·23**	
		68 29***	—	140 70***	8	7·97	0 67±0·05
		64·73***	−27·41	170·10***	7	7·04	
Geography	10	111·50***	101·60***	—	8	15 57*	
		74.06***	—	140 30***	8	11 12	0 65±0·08
		76·66***	23·42	115 40*	7	11 08	
Mathematics I	F	198·70***	—	—	3	6·37	
		149·30***	50 70	—	2	5 34	
		158·80***	—	39 70	2	5 23	
		179 80***	266·70**	−235 10	1	2·41	
	M	202·50***	—	—	3	8·54*	
		101·40***	104·60**	—	2	2 11	
		60·86**	—	147·40***	2	0·28	0·71±0·12
		61 72**	11·19	136 20	1	0 24	0·65±0 45
Biology	10	80·92***	81·70***	—	8	7·52	
		54·32***	—	107 40***	8	5 94	0·66±0 08
		59 90***	41 17	62 94	7	5·12	0 38±0·26

† Results are shown for females (F), males (M) and females, males and opposite-sex (10) where appropriate

* P<0 05 ** P<0 01 *** P<0·001

that all variation was due to individual environmental experiences failed badly and is omitted from the table. For all subjects (except Mathematics I) the same model was appropriate for both sexes, so models were fitted to all ten mean squares. For Mathematics I, models have been fitted separately to the four mean squares for females and the four mean squares for males.

For ASAT Arts, English and Geography, a purely environmental (E_1E_2) model fails to account for the data while the simple genetic model (E_1V_A) gives a good fit. For these three variables no further reductions in chi-square are obtained with the addition of an E_2 parameter to the model.

Although the E_1V_A model provides an adequate description of the ASAT Science data, addition of the E_2 parameter results in significant estimates of all three sources of variation indicating that the $E_1E_2V_A$ model is most appropriate. This is despite the fact that the improvement in chi-square over the E_1V_A model is only of marginal significance ($\chi_1^2 = 3.32$, $0.05 < P < 0.10$).

For Mathematics I, the E_1 model is able to describe adequately the data in females, indicating that there is no significant between families variation to be accounted for and that all variance is due to individual environment experiences. Although in males the data are consistent with both an E_1E_2 or E_1V_A model, the genetical model gives a better fit. When all three parameters are fitted, non-significant estimates are obtained for both $\hat{E}_2$ and $\hat{V}_A$, and $\hat{E}_2$ accounts for a trivial proportion of the variance (about 5 per cent).

For Biology, as with Mathematics I in males, there is little to choose between the E_1E_2 and E_1V_A models, although the latter gives a slightly better fit. Again, where all three parameters are fitted, non-significant estimates are obtained for $\hat{E}_2$ and $\hat{V}_A$, although $\hat{E}_2$ accounts for about 25 per cent of the variation.

The causes of covariation between ASAT Arts and Science scores

We have detected genetic variation for performance in all subjects (with the possible exception of Mathematics in females) and we may ask whether each subject is probing the same variation in a different way or whether some gene action is specific to certain subjects. Since not all twins have sat examinations in the same subjects we cannot employ the most elegant methods to answer this question (Martin and Eaves, 1977; Martin *et al.*, 1984). However, all twins since 1976 have taken the ASAT Arts and Science tests. The correlation between scores for these two tests is 0.58 and we wish to know the basis of this covariation. The procedure is described in detail in Eaves *et al.*, (1978, pp. 301-304). The first step is to calculate the 4 × 4 covariance matrices between the Arts and Science scores for Twin 1 and Twin 2, one matrix for each of the five twin types. Models are then fitted to the covariance matrices in which the variation and covariation between Arts and Science scores is attributed to various combinations of the three major sources, E_1, E_2 and V_A. For an E_1E_2 model of variation and covariation a fit of $\chi_{44}^2 = 42.7$ was obtained while the corresponding E_1V_A model yielded $\chi_{44}^2 = 23.1$. When covariation and variation from all three sources was specified in the model, a fit of $\chi_{41}^2 = 18.2$ resulted. Although none of the models is formally rejected by the chi-square criterion we note that the three source model represents a large and significant improvement over the E_1E_2 model ($\chi_3^2 = 24.5$, $P < 0.001$) but not over the E_1V_A model ($\chi_3^2 = 4.9$, $0.10 < P < 0.20$). We conclude that genetic factors are an important cause of covariation between Arts and Science performance but that E_2 is not. Estimates for the variance components for Arts and Science are similar to those from the univariate model fitting in Table 7 and the components of covariance for the three source model are $V_A = 31.0 \pm 8.5$, $E_1 = 6.9 \pm 2.2$, $E_2 = 5.7 \pm 7.9$. These translate into genetic, E_1 and E_2 correlations of 0.73, 0.37 and 0.43 respectively.

DISCUSSION

Twins taking the Tertiary Admission Examination in Western Australia from 1975 to 1981 appear typical of the other candidates in their cohorts. There is no indication of substantial heterogeneity in the twin samples between years nor between zygosity groups after pooling over years. Greater concordance between MZ than between DZ twins in the choice of examination subjects suggests that there is a genetic component in this choice.

Our analysis of the causes of variation in examination performance demonstrates that a significant proportion of this variation in each subject (except perhaps Mathematics in females and Biology) must be due to genetic factors. A model in which we sought to explain all variation in terms of environmental influences unique to the individual (E_1) or shared with co-twins (E_2) was firmly rejected. A model invoking additive genetic variation, however, gave a good fit in each case and no further significant improvement in fit could be obtained by allowing for shared environmental variation as well. Our results are very similar to those of other studies of examination performance on younger twin samples in South Australia (Martin, 1975) and Brazil (Salzano and Rao, 1976).

None of the above is to suggest that the *only* source of variation between pairs is due to the segregation of additive genes, but simply that our study, although bigger than most, does not have the statistical power to detect any other source. It has been shown that very much larger samples of twins than those available here are needed to reliably detect V_A and E_2 when both sources are making substantial contributions to trait variation (Martin *et al.*, 1978). The inability to resolve the causes of variation in Mathematics and Biology is largely due to the very small numbers of twins who were candidates for examination in those subjects.

A further complication in the interpretation of our results is that our twin sample is not typical of the population cohort as a whole but only of that portion of the cohort who were candidates for the TAE. This portion is only 34 per cent of the age cohort and we assume, in the absence of evidence to the contrary, that they represent the top 34 per cent from an underlying normal distribution of ability. Martin and Wilson (1982) have considered the bias that such truncate selection of twin samples will introduce to the calculation of genetic and environmental components of variation. Correlation coefficients estimated from such truncate samples will be biased downwards from their true population values, but if the trait has a genetic component then this bias will not be in the same proportion in MZ and DZ twins. In the present case let us take rough median values from our data of $r_{mz} = 0.70$ and $r_{dz} = 0.35$ which are consistent with a heritability of 0.70 and no E_2 variance. If we now interpolate in Figure 1 of Martin and Wilson (1982) using 66 per cent (100-34) sample truncation, we would obtain approximate estimates for the underlying ability trait in the population of $R_{mz} = 0.87$ and $R_{dz} = 0.65$. These putative population correlations would now be consistent with a random mating heritability of 0.44 and an E_2 component of 0.43.

Before assuming that this correction for truncate sampling has reinstated shared family environment to an important place in the aetiology of individual differences in educational achievement, we must remember that what we have estimated above as "E_2" also contains any additional genetic variance arising from assortative mating. There is certainly detectable covariation between spouses for educational achievement, estimates of the marital correlation ranging from about 0.3 to 0.7 (Jencks, 1972; Heath *et al.*, 1984). Even under moderate levels of assortative mating the extra additive genetic variation generated by assortative mating is considerable (Eaves *et al.*, 1978). Taking a median correlation between husbands and wives for educational attainment of 0.4, it can be calculated (see Martin, 1978) that more than half the variance attributed to "E_2" is genetic variance arising from assortative mating.

Taking all the above considerations into account, we would conclude that the total proportion of variance due to genetic factors is in the range of 60-70 per cent and that shared environmental influences (home background, shared schooling etc.) account for no more than 20 per cent of the total variance in examination performance. These proportions are very similar to the best estimates for the causes of variation in IQ (Fulker and Eysenck, 1979) and suggest that a distinction between IQ tests and tests of scholastic achievement on the basis of their heritabilities (e.g. Husén, 1960; Jensen, 1969) is not justified (cf. Willerman *et al.*, 1977).

We have been able to show that much of the covariation between ASAT Arts and Science scores is genetic in origin and that environmental factors contribute little to their correlation. The fact that the genetic correlation is only 0.73 indicates that, in addition to the common genetic factor underlying covariation, there are some differences in the genetic effects on the two traits. There is thus some support for the view that it is principally genetic factors which determine general ability and mainly (but not exclusively) environmental factors which channel it into proficiency in a particular subject area such as arts or science.

The finding of both common and specific genetic variation for performance in different tests is in line with the results from multivariate analyses of other twin studies of multiple abilities (Martin, 1975; Martin and Eaves, 1977; Martin *et al.*, 1984). A feature of the latter two studies, however, was that the "E_2" correlations between variables were all near unity. Eaves *et al.*, (1984) have argued that this is a consequence of assortative mating and/or cultural transmission operating on a latent variable which is a linear combination of the measured ability variables. In the present study we suggest that the low value of the "E_2" correlation is due to the combined effects of truncate sampling and the low power available to distinguish between additive genetic variation and other sources of variation between families.

The individual environmental correlation of 0.37 is small but the fact that it is greater than zero suggests that not all E_1 variance can be written off as "error". Heritabilities corrected for unreliable "error" variance will, of course, be higher than those quoted above but it is not clear what is gained by this practice, especially since it has been shown that "unreliability" itself may have a genetic component (Eaves and Eysenck, 1976). To the extent that an individual's career prospects are determined by his performance in a single examination, one is interested only in the causes of variation in that single measurement. However, to the extent that examination performance is an imperfect measure of an underlying trait on which, for instance, mate selection is based, heritability of the latent variable may be of more interest.

ACKNOWLEDGMENTS. — We would like to thank the Board of Management of the Tertiary Institutions Services Centre who gave us access to the ASAT and TAE results and Dr. N. A. Goodchild of the Biometrics Unit, University of Western Australia for his supervision of C. E. Gill. We are grateful to Dr. L. J. Eaves and Dr. A. C. Heath for helpful comments on the manuscript. Most of all we are indebted to the twins, without whose co-operation this work would not have been possible.

Correspondence should be addressed to Dr. N. G. Martin, Department of Human Genetics, Medical College of Virginia, Virginia Commonwealth University, PO Box 33, MCV Station, Richmond, Virginia 23298, USA.

REFERENCES

CEDERLÖF, R., FRIBERG, L., JONSSON, E., and KAIJ, L (1961). Studies on similarity diagnosis in twins with the aid of mailed questionnaires. *Acta Genetica et Statistica Medica*, 11, 338-362.

CLARK, P., JARDINE, R., MARTIN, N. G., STARK, A. E., and WALSH, R. J. (1980). Sex differences in the inheritance of some anthropometric characters in twins. *Acta Geneticae Medicae et Gemellologiae*, 29, 171-192.

Eaves, L. J. (1975). Testing models for variation in intelligence. *Heredity*, 34, 132-136.

Eaves, L. J., and Eysenck, H. J. (1975). The nature of extraversion: a genetical analysis. *J Person. soc. Psychol.*, 32, 102-112.

Eaves, L J., and Eysenck, H. J. (1976). Genetic and environmental components of inconsistency and unrepeatability in twins' responses to a Neuroticism questionnaire. *Behav. Genetics*, 6, 145-160.

Eaves, L. J , Heath, A. C., and Martin, N. G. (1984). A note on the generalised effects of assortative mating. *Behav. Genetics*, 14, 371-376.

Eaves, L. J., Last, K. A., Young, P. A., and Martin, N. G. (1978) Model-fitting approaches to the analysis of human behaviour. *Heredity*, 41, 249-320.

Fulker, D. W., and Eysenck, H. J. (1979). Nature and nurture: heredity. In Eysenck, H. J. (Ed), *The Structure and Measurement of Intelligence* New York: Springer Verlag.

Heath, A C., Berg, K., Eaves, L J., Solaas, M. H., Sundet, J., Nance, W E., Corey, L. A., and Magnus, P (1985) No decline in assortative mating for educational level *Behav. Genetics*, 15, 349-370.

Horn, J. M., Loehlin, J. C., and Willerman, L. (1982). Aspects of the inheritance of intellectual abilities. *Behav. Genetics*, 12, 479-516.

Husén, T. (1960). Abilities of twins. *Scand. J. Psychol.*, 1, 125-135.

Jencks, C. *et al.* (1972). *Inequality: A Reassessment of the Effect of Family and Schooling in America.* Harmondsworth: Penguin.

Jensen, A. R. (1969). How much can we boost IQ and scholastic achievement? *Harvard educ Rev*, 39, 1-123.

Jinks, J. L , and Fulker, D. W. (1970). Comparison of the biometrical genetical, MAVA and classical approaches to the analysis of human behaviour. *Psychol. Bull.*, 73, 311-349.

Kasriel, J., and Eaves, L. (1976). The zygosity of twins: further evidence on the agreement between diagnosis by blood groups and written questionnaires. *J biosoc. Sci.*, 8, 263-266

Loehlin, J. C., and Nichols, R C. (1976) *Heredity, Environment and Personality.* Austin. University of Texas Press.

Martin, N. G. (1975). The inheritance of scholastic abilities in a sample of twins II. Genetical analysis of examination results. *Annals of Human Genetics*, 39, 219-228.

Martin, N. G. (1978). Genetics of sexual and social attitudes in twins. In *Twin Research: Psychology and Methodology*. New York: Alan R. Liss.

Martin, N. G., and Eaves, L. J (1977). The genetical analysis of covariance structure. *Heredity*, 38, 79-95

Martin, N G., Eaves, L. J , Kearsey, M. J., and Davies, P. (1978). The power of the classical twin study. *Heredity*, 40, 97-116.

Martin, N. G., Jardine, R., and Eaves, L. J. (1984). Is there only one set of genes for different abilities? A reanalysis of the NMSQT data. *Behav Genetics*, 14, 355-370.

Martin, N G., and Martin, P. G. (1975). The inheritance of scholastic abilities in a sample of twins I. Ascertainment of the sample and diagnosis of zygosity. *Annals of Human Genetics*, 39, 213-218.

Martin, N. G., and Wilson, S. R. (1982). Bias in the estimation of heritability from truncated samples of twins. *Behav. Genetics*, 12, 467-472.

Nichols, R. C., and Bilbro, W C. (1966). The diagnosis of twin zygosity *Acta Genetica et Statistica Medica*, 16, 265-275.

Rao, D. C., Morton, N. E., Elston, R. C., and Yee, S. (1977). Causal analysis of academic performance. *Behav. Genetics*, 7, 147-159.

Rice, J., Cloninger, C. R., and Reich, T. (1980). Analysis of behavioral traits in the presence of cultural transmission and assortative mating. *Behav. Genetics*, 10, 73-92.

Salzano, F. M., and Rao, D. C. (1976). Path analysis of aptitude, personality and achievement scores in Brazilian twins. *Behav. Genetics*, 4, 461-466.

Willerman, L., Horn, J. M., and Loehlin, J. C. (1977). The aptitude-achievement test distinction: A study of unrelated children reared together. *Behav. Genetics*, 7, 465-470.

(*Manuscript received* 21*st March*, 1984)

Br. J. educ. Psychol., 55, 251-263, 1985

THE DETERMINANTS OF TEACHERS' ATTITUDES TO INTEGRATING THE INTELLECTUALLY HANDICAPPED

By DAVID THOMAS
(*Flinders University of South Australia*)

SUMMARY. The main purpose of this study was to identify those factors which serve as the best predictors of teachers' attitudes to integrating the intellectually handicapped. The factors that were investigated were selected on the basis of a review of the relevant research literature, a consideration of psychological and sociological theories of attitude formation and interviews with teachers and educational psychologists.

Teachers' attitudes to integration were assessed by a Q-sort, the content of which was derived from a review of the integration literature and discussions with teachers. The sample consisted of 550 primary and secondary teachers, special class and special school teachers, withdrawal teachers, special education advisers and school principals in Devon and Tucson, Arizona.

The identification of interaction among the independent variables was the main feature of the data analysis. The attitude of the "contact" special educator, conservatism as a dimension of personality, type of teaching, the tradition of special educational organisation in the area and the teacher's perception of her competence in selecting appropriate methods were shown to be significant factors in teachers' support for or opposition to integration.

INTRODUCTION

THE justification for this study lies in the assumption that the attitudes of teachers to integration have a major effect on its successful implementation (Mathey, 1977; Warnock, 1978). These attitudes have taken on a special importance now that integration has become an integral part of special educational services throughout the English world. In examining teachers' attitudes to integration, this study focuses on the relative significance of those factors that may be related to attitudinal differences in the teaching profession, in the belief that the identification of these factors would have implications for those initiating integration practices.

The significance of the study lies in the first place in its attempt to find and measure interaction effects between factors associated with differences in attitudes to integration, thus avoiding the dubious assumption of the 'independence' of the independent variables. Secondly, the study is set within a theoretical framework which incorporates explanations of attitude formation drawn from the broad spectrum of the social sciences as well as from research into teachers' mainstreaming attitudes. Data for the study were collected in two locations, Tucson and Devon, where there are markedly different traditions of special education for the intellectually handicapped. This enabled a test of the theory that attitudes are conditioned over time by tradition and practice (Bem, 1970).

Until recently, teachers' attitudes to integration have received little attention from academic researchers and earlier research has been, in the opinion of Watson and Hewett (1976), of a superficial nature. With few exceptions, notably Mandell (1976) and Stephens and Braun (1980), there has been a tendency to concentrate on a few readily verifiable factors such as length of teaching and age. Scant attention has been given to theories of attitude formation and in particular to institutional and personality factors. It is perhaps not surprising that in the author's review of the literature (Thomas, 1982) the highest explained variance in those studies quoting relevant statistics, is that of Mandell (1976) who accounted for just 23 per cent of the variance.

To provide a more comprehensive coverage, the factors chosen for this study were taken from a review of the related special education literature including unpublished doctoral dissertations, a consideration of theories of attitude formation in social psychology and sociology, the literature on personality, and interviews with individual teachers and educational psychologists in which they were asked to estimate which factors were associated with support for or opposition to integration. As a result, 39 variables of a societal, institutional, personal or personality type were included in the investigation.

The study was limited to the intellectually handicapped because it is the acceptance of these children that seems to raise the most sensitive issues for the ordinary teacher and provokes the most disagreement about the wisdom of integration. The teachers in the study were more familiar with the terms ESN(M) or EMR, so these were used in their interviews. 'The intellectually handicapped' is used here as a convenient synonym and is preferred for the sake of brevity to 'children with moderate learning difficulties'.

METHOD

Measuring the dependent variable

The preference for developing a new instrument stemmed from dissatisfaction with existing instruments and from a belief that accurate judgments of teachers' integration attitudes could not be made only through interviews or questionnaires. Interviewer bias, acquiescence response set, the social desirability factor, and cynicism towards questionnaires among some teachers were important reasons for rejecting the more usual techniques. Equally significant was the fact that the forced choice Q-sort ensures that respondents make comparative judgments between issues as well as reflecting on each issue separately, which would appear to be an essential feature in forming an attitude to a complex issue.

Thus the Q-sort technique was used to measure attitudes to integrating the intellectually handicapped. This was of the forced choice, unstructured, unimodal type and entailed a modification of the standard Q-sort technique, providing measures of the respondent's degree of opposition or support on 21 issues of integration.

The development of the instrument followed interviews with teachers which indicated that an attitude to integration took one of three forms:— (1) support or opposition varying according to the degree of integration; (2) a recognition of the difference between the educational and social development of the handicapped child in an integrated setting; or (3) a conceptualisation of the implications of integration for each of the participants involved.

Since this last conceptualisation was the most frequent, the Q-sort was designed to provide separate measures of whether the respondent thought integration was good for the handicapped child; good for other children in the class; whether it could be organised and accommodated within the ethos of the ordinary school; and finally whether the ordinary teacher could cope successfully with it. The respondent's Q-sort score, which could range from -68 to $+68$, was the total of these sections' scores.

The Q-sort, originally having 70 items, was reduced to 60 items and then to 42 items following a pilot study which included a test-retest measure of reliability (Pearson's $r = 0{\cdot}967$). An item analysis and a comparison of independent judges' ratings of essays on integration written by teachers and those teachers' responses to the 60 item Q-sort (Pearson's $r = 0.868$) provided the basis for checking the validity of the instrument (Thomas, 1982).

Measuring the independent variables

Since the main purpose was to explain differences in teachers' attitudes to integrating the intellectually handicapped, it was necessary to collect information about sources of potential influence on such attitudes for each participant in the study. The sources of influence were categorised as societal and institutional factors, personal experience, the notion of the significant other, non-experiential factors and personality.

Societal factors

The powerful impact of social norms on individual behaviour, beliefs and attitudes is a basic tenet of sociology (Silverman, 1970) and their influence is thought to be inextricably linked with the culture in which the individual grows up (Kelvin, 1970). It follows that teachers' attitudes to integrating the handicapped could be expected to be affected by the concern for the handicapped of the larger society. In the case of British and American cultures, the consensus of opinion suggests that people have become increasingly tolerant of the handicapped. In the absence of evidence to the contrary it cannot be assumed that this growth of empathy is appreciably greater in one or other of these countries.

Nevertheless the attitudes of a group within a community rather than society at large may be fashioned by local tradition. More precisely, a tradition of integrating the intellectually handicapped can be expected to have different implications for teachers' attitudes to integration than a tradition of special classes and special schools. Such a prediction would be consistent with Bem's (1970) self-perception theory which drew attention to the attitudinal effects of institutionalised and routine behaviour.

Notwithstanding the ensuing disagreement among social psychologists over Bem's theory, its implications were clear for this study. The possibility of different traditions of special education influencing teachers' attitudes to integration had to be taken into account in siting the study. On the basis of Birch's (1974) review of mainstreaming practices for the intellectually handicapped in USA, Tucson Unified School District with its long history of integrating these pupils and, on the basis of the writer's experience, Devon with its tradition of special classes and special schools for the ESN(M) were chosen as appropriate locations. If a tradition of integration did incline teachers to support integrating the intellectually handicapped, Tucson teachers would reflect this in higher positive scores on the Q-sort than their Devon counterparts.

Institutional factors

Within the concept of the school as a formal organisation influencing professional communications and attitudes, it was accepted that teachers' attitudes to integration could be affected by some physical aspects of the school and by some of its organisational structures. Accordingly, certain aspects of the physical environment, namely school size (Barker and Gump, 1964), location (Havighurst, 1964) and open plan/traditional architecture (Saylor and Alexander, 1974), were included in this investigation.

With regard to the influence of organisational structures, a school's adoption of an integration policy immediately affords the policy a degree of legitimation in the minds of many teachers. This may be reinforced or confounded by the nature of formal leadership in the school and the headteacher's attitudes towards the adoption of an integration policy. Practices within the school, particularly allocation of students to classes, may also be discordant with the declared policy of the school district, leading teachers to interpret these practices in their school as indicating covert opposition to the official school district position. If compliance at the school level is a more powerful determinant of attitudes than Bem's notion of the effects of tradition, then the mainstreaming attitudes of teachers in 'discordant' schools would not on the whole reflect the official school district policy. Thus details were included regarding the current school policy on integration, the direction of any intended changes, the nature of special education services within the school (withdrawal groups, special classes, resource teachers, visiting advisers or any combination of these) and the extent of streaming in the school.

Role differentiation can also have a bearing on teachers' attitudes to specific issues. It has been suggested that divergent views and vested interests increase with greater differentiation within the organisation (Beck and Betz, 1975). On this basis teachers in secondary schools might be expected to be more resistant to integration than primary teachers. Since the effects on teachers' attitudes of any of these factors may vary in schools with an exceptional incidence of social disadvantage, details were collected about which schools in the sample were designated as educational priority schools.

Personal factors

The effect of experience and personal interaction on attitude formation has been strongly emphasised by social psychologists. Given the interaction that can take place between a teacher and her principal and more particularly between a teacher and a special educator who teaches or acts as an adviser in the school, the concept of the significant other seemed especially apposite in a study of teachers' attitudes to integration. It was, however, necessary to distinguish between a person who fills one of these roles and is potentially influential and one who is in reality influential (Sheriff *et al.*, 1965). In other words, a special educator may be a significant other to one teacher but not to another in the same school. Such a distinction was incorporated in this study to facilitate assessment of this concept. Equally important was the position of the special educator on integrating the intellectually handicapped and whether this was made known to the teacher.

In addition, the view that social roles strongly influence attitudes in line with an acceptance of the values associated with those specific roles (Kelvin, 1970) implies that teachers in different roles may develop different attitudes to integrating the intellectually handicapped. Special class teachers, special education advisers, principals and ordinary class teachers may vary in their acceptance of integration irrespective of organisational structures and this possibility was allowed for in the data analysis. Similarly, the receptivity of first year teachers to the organisational structures of the school (Secord and Backman, 1974) and the pressures on newcomers to conform to the values and practices of the school (Lortie, 1975) raise the question of whether the attitudes of those with little professional experience would be more consistent with the status of integration within the schools than the attitudes of their more experienced colleagues. Consequently length of teaching experience was added to the list of independent variables.

Several other potentially relevant personal factors were identified by previous studies of teachers' mainstreaming attitudes. Some of these related to the individual as a teacher, namely:—experience of teaching the handicapped (Overline, 1977); year level taught (Watson and Hewett, 1976); teacher-aide (Mandell, 1976; Weber, 1977); in-service training (Fiorentino, 1978); and teacher competence (Carey, 1978).

All of these were regarded as important factors to be included in the study, though teacher competence presented a problem of assessment. The subtle distinction, suggested by Murphy (1960), between actual competence and the teacher's perception of his own competence offered an acceptable solution in the light of Stephens and Braun's (1980) finding that perceived self-competence accounted for as much as 19 per cent of the variance in teachers' scores on a test of attitudes to integration.

Attention has also been drawn to family related variables as potential sources of influence. These include the marital and parental status of the teacher and the teacher as a parent of a handicapped child (Stephens and Braun, 1980). These factors and the widely researched non-experiential variables of sex and age of the teacher were incorporated in the study. The list of independent variables was extended by six South Australian teachers and eight British educational psychologists who were asked to consider which factors would serve as the best predictors of teachers' attitudes to integrating the intellectually handicapped. They cited class size; the extent of contact between the ordinary teacher and a special educator; time spent in team teaching; and any qualification in special education. Details were also included in the analysis of whether any member of the family was handicapped or worked with the handicapped; whether the respondent had had any recent close contact with the handicapped outside of the classroom; what the nature of previous teaching experience had been in terms of year levels taught and the emphasis of any recent in-service course on integration in terms of support or opposition to integration.

Personality

The last of the independent variables mentioned earlier was assessed by the Wilson-Patterson Attitude Inventory (WPAI) which gives separate measures of Conservatism-liberalism, Realism-idealism and four minor components, Ethnocentrism-intolerance, Anti-hedonism, Religion-puritanism and Militarism-punitiveness. The first two of these factors were considered particularly apposite. The concept of conservatism with its implications for the acceptance or rejection of change has obvious relevance to teachers' attitudes to integration. The construct of Realism-idealism must also be considered as a potential source of explanation of differences in teachers' integration attitudes. Conceivably the idealist would be more inclined to tolerate the introduction of the intellectually handicapped into ordinary classes.

Variables not included

Several factors cited in the literature or in the interviews were not incorporated in the main analysis. Traumatic experiences with the handicapped, for example, were omitted from the AID analysis because of the difficulties in allowing for possible exaggeration in teachers' accounts when directly questioned about experiences of this kind. Spontaneous comments about traumatic experiences in interviews, however, were recorded. Details of membership of a teachers' union, the religion and measured intelligence of each respondent were omitted for reasons of expediency.

Sampling

A two-stage sampling frame was developed in which schools were randomly selected with a minimum participation of three teachers and the principal from each school. Equal numbers of teachers in Tucson and Devon were sought to provide a fair test of the theory that traditions and policy affect attitudes over time. Since information on type of teaching was readily accessible, it was decided to seek equal numbers of special education teachers, high school teachers, lower primary and upper primary teachers in the two locations. It was also considered advisable to separate the group of special educators according to the nature of their work, with an equal number of special class teachers, withdrawal group teachers and special or remedial education advisers in the two locations.

It proved impossible to go beyond this and to incorporate other independent variables into the stratification of the sample. The actual sample closely matched the basic requirements with ten secondary schools and four special schools in the two areas participating in the study. The number of primary schools varied slightly, with 24 in Tucson and 27 in Devon. Table 1 gives further details of the sample.

Since the refusal rate was only 3.6 per cent no attempt was made to weight for non-response.

TABLE 1

TEACHERS AND PRINCIPALS IN THE SAMPLE

	Primary	Secondary	Principals	Special Class Teachers	Other Special Educators	Total
Tucson	100	65	30	24	48	267
Devon	101	71	31	27	53	283
	201	136	61	51	101	550

Data analysis

The importance of identifying and ordering interaction effects for a proper understanding of the relationship between independent and dependent variables in exploratory research is well documented (Blalock, 1968; Sonquist, 1970; Sonquist *et al.*, 1973). Ample evidence exists in educational research to support this view (Entwistle and Welsh, 1969; Ainley, 1976; Evans, 1976). In particular, several studies indicate the significance of interaction between several independent variables to teachers' attitudes to integrating the intellectually handicapped (Bishop, 1969; Watson and Hewett, 1976; Overline, 1977).

How then should interaction be measured? In defining interaction as the non-additivity of the effects of two or more independent variables on the dependent variable, Blalock (1968) has drawn attention to the limitations of the majority of multivariate techniques, with their focus on additivity in analysing data containing important interactions. While regression analysis can measure interaction effects where the number of variables is small, it is virtually impossible to compute all the necessary interaction terms in cases where there are many predictors (Westervelt, 1960). Covariance techniques which can detect interaction effects have a similar problem where there are several such effects among a large number of predictors (Sonquist, 1970).

What was required was a procedure which searched for interaction effects among those predictors which together account for most of the variance in the dependent variable without assuming additivity and linearity and locates interactions in that part of the sample affected by them. Automatic Interaction Detection Analysis (Aid 3) (Sonquist *et al.*, 1973) meets these requirements and was used in this study.

The AID algorithm is essentially the stepwise application of one-way Analysis of Variance (ANOVA) in a form that is unique. It seeks to partition the total sample into two non-overlapping groups so that their means account for more of the variation in the dependent variable than can be achieved by any other split. To do this, each independent variable must first be divided into two or more non-overlapping categories. ANOVA statistics are then calculated for every possible split in each predictor. The programme notes the division in each predictor which makes for the greatest reduction in error by attempting to maximise the between-sum of squares while reducing the within-groups variation to a minimum. The division which produces the greatest error reduction overall is considered the basis for splitting the total sample.

In this way the total sample is partitioned into two mutually exclusive groups by one of the independent variables at a critical point in its range. The two 'daughter' groups are then examined and split in the same way. The process continues to produce more differentiated and mutually exclusive subgroups in a branching structure, reflecting only those effects which serve to explain variance in the dependent variable.

RESULTS

As mentioned earlier, the main purpose of the study was to examine major sources of influence upon teachers' attitudes to integrating the intellectually handicapped. The following discussion is based on an AID analysis of the relationship between these factors, described earlier as potential predictors of attitudes to integration, and the scores on a Q-sort designed to measure degree of support or opposition to integrating the intellectually handicapped.

AID analysis

The AID analysis accounted for 30.3 per cent of the variance in the dependent variable, exceeding that of any of the studies mentioned in the author's review of the literature (Thomas, 1982). Since most teachers in the Devon and Tucson samples were opposed to integrating the intellectually handicapped as reflected in the overall mean Q-sort score of −10, the distribution of Q-sort scores was skewed (skewness coefficient of 0.419, range of scores +58 to −58). Consistent with the need to minimise skewness in the dependent variable for a proper examination of the relationship between the independent variables and the dependent variable (Sonquist, 1970), a logarithmic transformation of the latter was made, reducing the coefficient of skewness to −0.014, well within the limits of normality.

The best single predictor in terms of the independent variable that accounts for more of the variance in the dependent variable than any other was the perceived attitude to integration of the contact special educator (Fig. 1). This variable alone explained 7 per cent of the total variance. In other words, the latter's position on integrating the intellectually handicapped, as teachers perceive it, has a notable effect on teachers in the same school or, where the contact special educator is itinerant, in schools visited by him. In the context of attitude theory, this draws attention to the importance of the concept of 'the significant other' in the formation of attitudes.

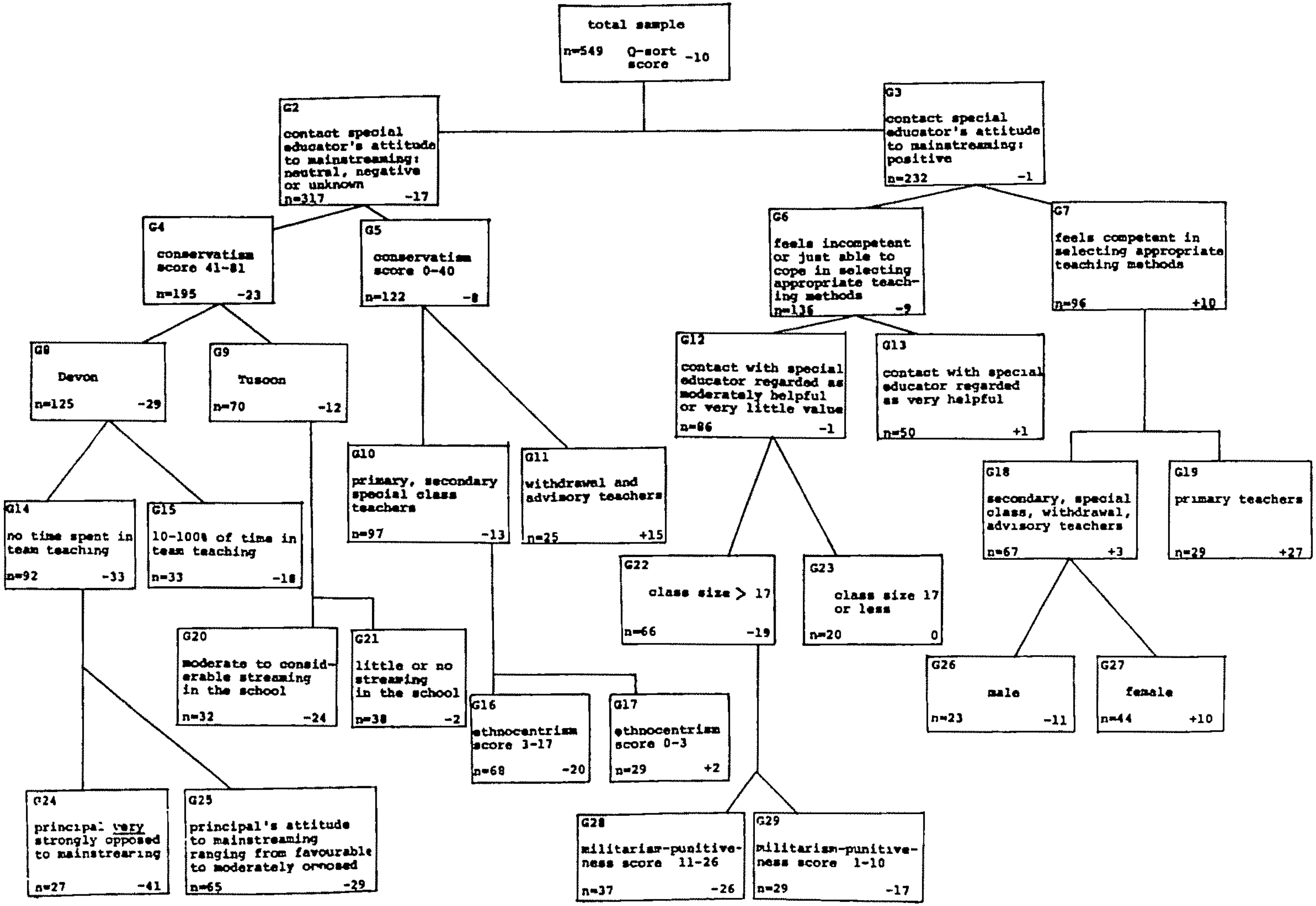
An examination of the split of the main sample into groups G2 and G3 (Fig. 1) and their mean Q-sort scores (−17 and −1 respectively) indicates that where the contact special educator's attitude is known to be positive, the effect is to reduce opposition to integrating the intellectually handicapped. Where this attitude is negative, the effect is the opposite. This is of course what we would expect. Less predictably, where his attitude is neutral or unknown the effect seems to be similar to a negative attitude. It would appear then that, if a special educator is in favour of integration but does not make known his support, then any effect on ordinary teachers is likely to be in the opposite direction from the one he prefers.

The personality trait of conservatism explains more of the variance than any of the remaining variables although its effect is mainly restricted to those teachers who believe their contact special educator is neutral or negative in his mainstreaming attitude (G2). Those group 2 teachers with Conservatism scores on the WPAI in the lowest 38 per cent are inclined to be less opposed to integrating the intellectually handicapped (G5; $\bar{x} = -8$) than other group 2 teachers with higher Conservatism scores (G4; $\bar{x} = -23$).

For those whose contact special educators were seen as having a positive attitude to mainstreaming (G3), confidence in selecting appropriate teaching methods and materials for the slow learner was the most discriminating factor. Those who perceived themselves as being competent (G7; $\bar{x} = +10$) were inclined to be appreciably more positive in their attitude to integrating the intellectually handicapped than those who felt just able to cope or were at a loss in selecting appropriate methods and materials (G6; $\bar{x} = -9$).

It is, of course, necessary to stress that confidence in selecting teaching methods is

FIGURE 1 — A I.D. ANALYSIS


total sample
n=549 Q-sort score -10
G2
contact special educator's attitude to mainstreaming: neutral, negative or unknown
n=317 -17
G3
contact special educator's attitude to mainstreaming: positive
n=232 -1
G4
conservatism score 41-81
n=195 -23
G5
conservatism score 0-40
n=122 -8
G6
feels incompetent or just able to cope in selecting appropriate teaching methods
n=136 -9
G7
feels competent in selecting appropriate teaching methods
n=96 +10
G8
Devon
n=125 -29
G9
Tuscon
n=70 -12
G10
primary, secondary special class teachers
n=97 -13
G11
withdrawal and advisory teachers
n=25 +15
G12
contact with special educator regarded as moderately helpful or very little value
n=86 -1
G13
contact with special educator regarded as very helpful
n=50 +1
G14
no time spent in team teaching
n=92 -33
G15
10-100% of time in team teaching
n=33 -18
G16
ethnocentrism score 3-17
n=68 -20
G17
ethnocentrism score 0-3
n=29 +2
G18
secondary, special class, withdrawal, advisory teachers
n=67 +3
G19
primary teachers
n=29 +27
G20
moderate to considerable streaming in the school
n=32 -24
G21
little or no streaming in the school
n=38 -2
G22
class size > 17
n=66 -19
G23
class size 17 or less
n=20 0
G24
principal very strongly opposed to mainstreaming
n=27 -41
G25
principal's attitude to mainstreaming ranging from favourable to moderately opposed
n=65 -29
G26
male
n=23 -11
G27
female
n=44 +10
G28
militarism-punitiveness score 11-26
n=37 -26
G29
militarism-punitiveness score 1-10
n=29 -17

Footnote Mean for G12 should read —15.

not of itself an important factor for the whole sample but attains its significance in its interaction with the contact special educator's attitude to integration. The combination of these two variables in inducing support for integration was especially pertinent to primary school teachers. Indeed this group had the highest positive Q-sort score (G19; $\bar{x} = +27$).

The splitting of group 6 (Fig. 1) suggests an important interaction between a lack of confidence in selecting appropriate methods for the intellectually handicapped and the quality of support offered by the contact special educator. The similarity in the means of G13 ($\bar{x} = +1$) and G3 ($\bar{x} = -1$) suggests that where the contact special educator proves 'very helpful' in the eyes of the regular teacher this compensates for that lack of confidence. On the other hand, where the regular class teacher finds the contact 'of little use' or 'helpful' rather than 'very helpful', then opposition to integration increases among those lacking confidence in their ability to teach the intellectually handicapped (G12; $\bar{x} = -15$). Comments in interviews after completion of the Q-sort suggested that 'helpful' was merely a euphemism for a relationship that was seen as marginally better than 'of little use' rather than being close to 'very helpful'. The interaction identified by the split on group 6 draws attention again to the concept of 'the significant other' as a force in attitude formation.

The group with the second highest positive score (G11; $\bar{x} = +15$) was moderately in favour of integrating the intellectually handicapped. This group was composed of special education advisers and withdrawal teachers who obtained low Conservatism scores. The wide discrepancy between the means of groups 10 and 11, means of -13 and $+15$ respectively, underlines the importance of type of teaching as a determinant of teachers' attitudes to mainstreaming. The split of G8 on the basis of team teaching is also worthy of comment. For several teachers, the realisation that a team teaching colleague shared their uncertainty about integration proved reassuring and served to reduce their opposition.

An examination of the left branch of the tree gives an indication of factors which in combination with each other militate most against acceptance of integration. A tradition of special classes and special schools in the area, limited opportunity for team teaching and a negative attitude from the principal to integrating the intellectually handicapped, along with moderate to high conservatism and an unsupportive attitude of the contact special educator, form a powerful set of circumstances with regard to negativism towards integration. These factors are reflected in group 24 whose mean of -41 indicates very strong opposition towards integrating pupils in this category of disability.

Perusal of the AID tree (Fig. 1) reveals interesting examples of the view of Blalock (1968) and Sonquist (1970) that interaction effects cannot be assumed to be constant over the entire sample. One of these is found in the division of G18 on the basis of the sex of the teacher. The greater tolerance of female teachers for integrating the intellectually handicapped was found only among a group of 67 non-primary teachers, confident in their ability to select appropriate teaching methods for the slow learner and in contact with a special educator who was perceived as supporting integration. In other words, sex of the teacher, like several other factors does have a bearing on teachers' integration attitudes but its effect is limited to a minority. A similar conclusion can be drawn with regard to the tradition of integration within the school district although it affects a much larger group.

Two provisos are necessary concerning the independent variables. First, a case analysis of G23 revealed that no member of this group was a regular class teacher, indicating that the variable class size served as a proxy variable for type of teaching. Thus the relationship of class size to teachers' mainstreaming attitudes had not been clarified by the AID analysis.

Secondly, the AID splitting process masks differences between teachers in Tucson and Devon as a whole and as definable groups. The strong opposition to integrating the intellectually handicapped in Devon is shown by the fact that only 19.5 per cent of teachers were in favour (Tucson 30.8 per cent). This was especially true of high school teachers and special class teachers where only 12 per cent and 6 per cent respectively indicated support (Tucson 17 per cent and 37 per cent respectively).

The independent variables as predictors of integration attitudes

Equally as interesting are those variables that had been identified by the literature or by teachers and psychologists as potentially important but which in fact proved not to be. This was true for 26 of the 39 independent variables. These included size of school, the organisation of special education within the school (whether it had a special class or withdrawal class, or both or neither), and whether the school had been designated as an 'educational priority' school. Included in the personal factors that made no significant contribution to the variance were the age and parental status of the teacher, whether they had a qualification in special education and the extent and duration of the contact of regular teachers with the handicapped in and outside the classroom. Surprisingly, no clear pattern of support for integration has emerged in the teachers who were parents of handicapped children.

TABLE 2

PERCENTAGE OF THE VARIANCE EXPLAINED BY VARIABLES IN THE AID ANALYSIS

Independent Variable	Percentage Variance
1. Attitude of contact special educator to mainstreaming	7·0
2. Conservatism and related personality traits	5·6
3. Type of teaching	5·4
4. Confidence in selecting appropriate teaching methods (for the slow learner)	3·3
5. Traditional policy of integration (location)	2·7
6 Policy of allocation of students to classes (mixed ability or streaming)	1·8
7. Quality of support offered by the contact special educator	1·4
8 Sex of the teacher	1·3
9 Team teaching	1·1
10 Principal's attitude to integration	0·7
Total variance	30·3%

The factors that featured as the relevant predictors are listed in Table 2, along with their contribution to the variance of the dependent variable. The percentage attributable to type of teaching is the sum total of the variance explained by its occurrence twice in the tree and by its proxy variable class size (for G23 all classes smaller than 17 pupils were special classes).

DISCUSSION

The balance of opinion among the teachers in this study is against the integration of the intellectually handicapped. This applies both to teachers in Tucson, with its mainstreaming traditions buttressed by official policy, federal legislation and generous financial support, and to Devon teachers who have experienced none of these pressures. Clearly, tradition and practical steps to implement policy cannot guarantee acceptance among those with day-to-day responsibility for its implementation. Yet the teaching profession in Tucson is significantly less opposed to integration than is its Devon counterpart, indicating that such a tradition, backed by extra resources and personnel, does influence attitudes over time. That influence appears to have its greatest effect on those who see themselves as charged with a special responsibility to uphold and extend that policy. In this study, this applied to Tucson special educators and, in particular, to its special education advisers.

Nevertheless, an examination of the various groups' Q-sort scores gives only qualified support for Barngrover's (1971) theory that support for integration correlates inversely with proximity to day-to-day responsibility for teaching the handicapped in integrated surroundings. According to this view, special education advisers would be expected to obtain higher positive Q-sort scores than head teachers, who in turn would have higher scores than withdrawal teachers, who would have higher scores than class teachers. While this was generally true for this sample, withdrawal teachers in Devon high schools were more positive overall than their head teachers and no different in their degree of support from Devon's peripatetic remedial education teachers. Similarly, the Q-sort of withdrawal teachers in Tucson primary schools were not significantly different from those of their primary school teachers.

The generality of the findings

It is claimed that the sampling techniques used in this study permit its results to be generalised beyond the sample to the teaching profession in Tucson and Devon. The generality of the results beyond those areas remains an open question. In so far as other school districts exist in the United States with a similarly long tradition of integrating the intellectually handicapped and with comparable special education resources to Tucson, it is feasible to predict results in keeping with those obtained by the Tucson sample. The same proposition can be extended from the Devon sample to local education authorities in Britain with similar traditions of segregated special education and equal resources. It is recognised, however, that this investigation has been exploratory in nature and any generalisations are inevitably tentative and restricted to attitudes to the intellectually handicapped.

A further constraint on the generality of these findings is the time factor, which Cronbach (1975) has suggested is an important but frequently ignored variable in educational and psychological research. In the context of this study it cannot be assumed that the pattern of support and opposition to integration found in 1979 would be much the same five or ten years later, though in the opinion of some of the participants, canvassed in September, 1984, little has changed meanwhile to modify teachers' attitudes. Nor can it be assumed, even in districts offering firm support for integration, that any attitudinal change across the teaching profession will inevitably reflect less opposition to mainstreaming the intellectually handicapped. Thom's (1972) catastrophe theory, which seeks to identify the crucial points at which minor variations in some predictor(s) trigger off large variations in a dependent variable, when applied to change in attitudes to social and political events (Isnard and Zeeman, 1976), has supported the belief that such change is not necessarily synonomous with greater tolerance or progress.

Attitude change

Evidence in this study suggests that part of the problem of attitude change lies in the communication gap between the teaching profession and those outside the schools who promoted and continue to support the concept of integration. The response of teachers to the Q-sort indicated, for example, that the notion of modelling, which implies that the adjustment of the handicapped child will be facilitated by the company of the non-handicapped, has not become part of the social stock of knowledge of regular class teachers. Yet it has been an important assumption among the advocates of integration. The same appears to be true of the self-fulfilling prophecy in that teachers in this study generally rejected the notion that the intellectually handicapped would make less progress in the special class. Special education advisers at school district level would seem to be key personnel in bridging this communication gap and thus could serve a vital role in laying the foundation for a greater acceptance of integration among the teaching profession. Particularly important is the need for special education advisers to consider ways in which they can sensitise themselves to the problems facing schools where opposition to integration remains high.

The difficulties in initiating attitude change are admittedly profound but the five schools project in Northern Colorado (Kreinberg and Chow, 1973) and more recent examples in Britain (e.g., Hallmark, 1983) indicate that change is possible. The former project underlined the significance of the active involvement of the school principal in implementing integration and the AID analysis in this study drew attention to the contribution that the contact special educator can make in encouraging support for it. The latter can exert considerable influence by his advocacy and by close support of those class teachers whose efforts to help the handicapped are impaired by a lack of confidence in their own ability.

Acknowledgments. — Thanks are due to Mr. Jocelyn Owen, C.E.O. for Devon and Mrs. Barbara Prentiss and Dr. Mary Meredith of Tucson Unified School district for their permission and co-operation in collecting the data; to Mr. Tong Ly and Dr. Mary Rose, Coombs Computing Centre, Australian National University, Canberra; and to my colleagues at Flinders University, Dr. M. Lawson, Dr. P. OBrien and Professor J. Darroch for their constructive suggestions at different stages of this study.

REFERENCES

Ainley, J. (1976). The Australian Science Facilities Programme: a study of its impact on schools. Paper presented at the 47th ANZAAS Congress, Hobart, May, 1976.

Barker, R G., and Gump, P V. (1964). *Big School, Small School.* Stanford, Calif. Stanford.

Barngrover, E. (1971). A study of educators' preferences in special education programmes. *Except. Children*, 37, 754-755.

Beck, E. M, and Betz, M. (1975). A comparative analysis of organisational conflict in schools. *Sociol. Educ*, 48 (Winter), 59-74.

Bem, D. J. (1970). *Beliefs, Attitudes and Human Affairs.* California: Brooks, Cole Publishing Co.

Bentler, P M., and Speckart, G (1979). Models of attitude-behavior relations *Psychol Rev.*, 86 (5), 452-464.

Birch, J. W. (1974). *Mainstreaming. Educable Mentally Retarded Children in Regular Classes.* Reston, Va: Council for Exceptional Children.

Bishop, S (1969). Attitudes toward the handicapped as related to demographic variables and past experience. Honours paper for the elementary laboratory course, University of Michigan

Blalock, H. M (1968). Theory building and causal inferences. In Blalock, H. M., and Blalock, A. B., *Methodology in Social Research.* New York: McGraw Hill.

Carey, J. A. (1978) The inter-relationship between teacher skill, knowledge, attitude toward, knowledge of and behavioral interactions with handicapped children in the regular classroom. Ph D. thesis, University of Connecticut. Ann Arbor, Michigan University microfilms.

Cronbach, L. J. (1975). Beyond the two disciplines of scientific psychology. *Am. Psychol.*, 10, 116-127.

Entwistle, N. J., and Welsh, J. (1969). Correlates of school attainment at different ability levels. *Br. J. educ Psychol.*, 39, 57-63.

Evans, R (1976). The prediction of educational handicap. *Educ. Res*, 19, 57-68.

Fiorentino, M (1978). A study of the effects of short term in-service programmes in regular classroom teachers' attitudes towards and knowledge of mainstreaming. Ph.D. thesis, Massachusetts University.

Fishbein, M, and Ajzen, I. (1975). *Belief, Attitude, Intention and Behaviour: An Introduction to Theory and Research.* Reading, Mass.. Addison-Wesley; York: Academic Press.

Hallmark, N. (1983). A Support Service to Primary Schools. In Booth, T., and Potts, P. (Eds.), *Integrating Special Education.* Oxford: Basil Blackwell.

Havighurst, R. J. (1964). *The Public Schools of Chicago.* Chicago: The Board of Education.

Isnard, C. A., and Zeeman, E. C. (1976). Some models from catastrophe theory in the social sciences. In Collins, L. (Ed.), *The Use of Models in the Social Sciences.* London: Tavistock Publications.

Kelvin, P (1970). *The Bases of Social Behaviour An Approach in Terms of Order and Value.* London: Holt, Rinehart and Winston.

KREINBERG, N., and CHOW, S. (1973). Configurations of change: the integration of mildly handicapped children into the regular classroom. Report of US Dept. of Health, Education and Welfare, National Institute of Education, National Center for Educational Communication.

LORTIE, D. C. (1975). *Schoolteacher: A Sociological Study*. Chicago: University of Chicago Press

MANDELL, C. J. (1976). Factors related to the regular teachers' attitudes towards mainstreaming mildly educationally handicapped children into the regular classroom. Ph.D thesis, American University, Washington, D.C.

MATHEY, J. P. (1977). The effect of an in-service training workshop for regular classroom teachers on their attitudes towards handicapped children. Unpublished Ph.D. thesis, University of Michigan.

MURPHY, A. T. (1960). Attitudes of educators toward the visually handicapped. *Sight Saving Review*, 30 (3), 157-161.

OVERLINE, H M. (1977). Mainstreaming — making it happen. Hayward, California State University, Department of Educational Psychology. ERIC document 149 514.

SAYLOR, J. G., and ALEXANDER, W. M. (1974). *Planning Curriculum for Schools*. New York: Holt, Rinehart and Winston.

SECORD, P F., and BACKMAN, C. W. (1974). *Social Psychology (2nd edit)*. New York: McGraw-Hill.

SHERIF, C W., SHERIF, M. and NEBERGALL, R. E. (1965). *Attitude and Attitude Change. The Social Judgement-Involvement Approach*. Philadelphia. Saunders.

SILVERMAN, D. (1970). *The Theory of Organisations· A Sociological Framework*. London. Heinemann

SONQUIST, J. A. (1970). *Multivariate Model Building· The Validation of a Search Strategy*. Ann Arbor, Michigan: Institute for Social Research, University of Michigan.

SONQUIST, J. A , BAKER, E. L., and MORGAN, J. N. (1973). *Searching for Structure*. Ann Arbor, Michigan· University of Michigan Press.

STEPHENS, T. M., and BRAUN, B L. (1980). Measures of regular classroom teachers' attitudes toward handicapped children *Except. Children*, 46, 4.

THOM, R. (1972). Stabilite structurelle et morphogenese, cited in ZEEMAN, E C. (Ed.) *Catastrophe Theory—Selected Papers*. Reading, Mass.: Addison-Wesley.

THOMAS, D. (1982). Teachers' attitudes to mainstreaming the intellectually handicapped. Unpublished doctoral thesis, Flinders University of South Australia.

WARNOCK, H. M. (Chairwoman) (1978). *Special Educational Needs*. Report of the Committee of Enquiry into the Education of Handicapped Children and Young People London: HMSO.

WATSON, P. C., and HEWETT, F. M. (1976). The learning handicapped Integration Inventory: assessing regular classroom teachers' perception of the effect of integrating mildly handicapped children. Technical report SERP, Graduate School of Education, UCLA.

WEBER, M B. (1977). Attitudes of public school teachers towards mainstreaming the mentally retarded. Paper read at Georgia Association of School Psychologists' Council for Exceptional Children, Atlanta. ERIC document ED 150 817.

WESTERVELT, F H (1960). Automatic system simulation programming, cited in SONQUIST, J A., *Multivariate Model Building*. Ann Arbor, Michigan: University of Michigan College of Engineering

WILSON, G. D. (1975) *Manual for the Wilson-Patterson Attitude Inventory*. Windsor: NFER.

(*Manuscript received 20th June*, 1985)

Br. J. educ. Psychol., 55, 264-272, 1985

THE NATURE OF REVERSIBLE LETTER CONFUSIONS IN DYSLEXIC AND NORMAL READERS: MISPERCEPTION OR MISLABELLING?

By PAMELA BIGSBY
(*University of East Anglia*)

SUMMARY. An experiment was carried out to examine the extent and locus of the confusions which reversible letters have been observed to produce in the reading behaviour of dyslexic children. Using an extension of the Posner paradigm to collect speed and error data, the processing of reversible letters by dyslexic children was compared with that of both chronological and reading age controls. Results suggested that reversible letter confusions are errors of labelling rather than perception, and that dyslexic readers differ from normal readers of the same chronological age not in their perception of reversible letters but in their processing of them, within the name code. It was proposed that remedial programmes should concentrate on letter to sound mappings, particularly of reversible letters, rather than on largely unproductive programmes aimed at improving perception.

INTRODUCTION

THE reading process has recently been represented by visual information processing models. According to these models the information obtained from brief visual stimuli, such as a single eye fixation in reading or a tachistoscopic exposure, is represented in different codes at different times after stimulus offset (Haber, 1969). These codes are described as a pre-processed iconic store, a visual code which represents the visual characteristics of a letter, and a name code which represents its linguistic characteristics. Dyslexia had been increasingly regarded as an information processing deficiency (Stanley and Hall, 1973; Jorm, 1979; Ellis and Miles, 1981), and a review of the literature reveals a trend away from explanations which involve a visual code deficit and towards those which implicate a name code dysfunction (Katz and Wicklund, 1971; Denckla and Rudel, 1976; Jorm, 1979; Ellis, 1981; Legein and Bouma, 1981).

Originally, reversible letters (those which have the same form but different orientation, i.e., *b*, *d*, *p*, *q*) were thought to be literally perceived wrongly by dyslexic readers, either through a lack of suppression of the mirror-image produced by the alternate hemisphere of the brain (Orton, 1925) or through actual 'misperception' based on incomplete visual information being obtained from the stimulus (Lyle and Goyen, 1975). However, Vellutino *et al.* (1972) found that dyslexic children correctly reproduced graphically many words on which they made numerous positional errors in oral reading (e.g., din/bin; cob/cod; lion/loin) and suggested that so-called perceptual inaccuracies, such as reversals, were a dysfunction of the verbal identification system and not a distortion in perception. Ellis (1981) tachistoscopically presented dyslexic children with letter-like forms which included reversals and rotations. He concluded that the dyslexic group showed no impairment in terms of either the speed or the accuracy with which they could judge two letters to be the same or different on the basis of their physical characteristics though his results did lend some support to the notion that "the dyslexics may be slightly less accurate" (Ellis, 1981, p. 209).

Liberman *et al.* (1971) found that poor readers performed well on a letter matching task involving reversible letters and concluded that poor readers show no impairment in the processing of isolated reversible letters but that "the problem of reversible letters is specific to words". This conclusion would appear to be premature, however, as the task only explored the matching of physical characteristics. There is, therefore, a need for an experimental study of reversible letter processing which would consider both the physical and the linguistic characteristics of these stimuli.

Posner and Mitchell (1967) devised an experimental technique whereby visual and name code latencies could be measured. Subjects were asked to decide, as rapidly as possible, whether two simultaneously presented letters were the same or different. Response latencies for physically identical stimuli (e.g., *AA*), were found to be significantly shorter (70-100 msec) than those for stimuli which were the same in name only (e.g. *Aa*). They suggested that the physically identical match was not influenced by the naming of the letters and response latency here was a measure of visual code production (and hence presumably some measure of the act of 'perception'). Alternatively, the response of 'same' to stimuli such as *Aa* must include name coding and response latency here was taken to be a measure of the time needed to produce a name code. This difference in response latencies for visual and name matches has come to be known as the Posner effect.

It is still a moot point whether the response of 'different' to physically different stimuli, such as *sr*, operates within the visual or name code. Posner and Mitchell (1967) found that a 'different' response to a physically different stimuli pair, such as *AB*, was about 70-100 ms faster under a physically identical instruction (response of same required only if stimuli physically identical) than it was under a name identity instruction. Thus the same stimulus response combination (*AB*/different) gave differing RTs depending on the instruction used to define 'same'. In Posner and Mitchell's view, therefore, a response of 'different' to stimuli such as *sr*, in the present experiment, would involve a name match. Alternatively, Ellis (1981) suggests that different responses to visually different upper case letters are warranted without recourse to name code comparisons, as the physical characteristics of the large upper case second letter of the same case pairs are, as a class, so distinctly different from those of lower case second letters of mixed case letter pairs. Bigsby (1984) likewise holds that the response of 'different' to physically different letter pairs can be reached within the visual code. She suggests an intermediate categorisation stage, operating within the visual code, which classifies all non-identical stimuli as 'same case' or 'mixed case'. If the non-identical stimuli are 'same case' then a response of 'different' can be given without recourse to a name code. This is consistent with Hawkins' (1969) findings that "comparisons terminate upon the detection of information sufficient for a correct response".

This study, operating within the Posner paradigm and assuming that physically different matches operate within the visual code, was intended to examine whether the mistakes made by dyslexic children, during the processing of reversible letters, lay within the visual code, the name code or both of these. If mistakes and longer response latencies were found to be located largely in the visual code then misperception would be implicated. Alternatively, if the majority of errors and longer response latencies occurred mainly within the name code then a linguistic dysfunction, or mislabelling, would be implied.

METHOD

Subjects

Three groups of ten children, each of nine boys and one girl, were tested. The dyslexic group (DYS) attended a Specific Learning Difficulties unit, on a part-time basis. All the DYS children chosen satisfied the following conditions: (*a*) Reading Age as measured by the Burt Rearranged Test was at least 1·5 years behind chronological age; (*b*) there was no evidence of any gross behavioural problem or organic disorder; (*c*) performance or verbal IQ, as measured on the WISC test, was greater than 90 (in fact in all but one case it was greater than 100).

Two control groups, chosen from city schools, were also tested. A chronological age control group (CAC), age range 9·3-11·5 years, were all of average or above reading age, as measured by the Spar Reading Test. All subjects were of average or above intellectual ability, according to teacher judgment.

A reading age control group (RAC) was matched with the DYS group for reading age, using the Burt Rearranged Test (range 5·6-9·6 years). All subjects were of average or above intellectual ability, according to teacher judgment. Whilst this method of assessment introduces some measure of error, this was not deemed to be crucial. It was not possible, in this instance, to employ a uniform measuring instrument across groups.

TABLE 1

MEANS AND STANDARD DEVIATIONS FOR THE THREE EXPERIMENTAL GROUPS

	Dyslexic	CAC	RAC
Chronological Age	10·3 (0·8)	10·3 (0·8)	7·2 (0·8)
Reading Age	7·6 (1·1)	11·2 (1·1)	7·6 (1·1)

Experimental design

Out of the 15 conditions used in the original experiment, eight are relevant here. Each condition consisted of pairs of letters. Figure 1 shows the letter pairs and conditions used.

Lower case letter pairs were included as the reading process usually involves lower case letters. Name different and physically different letter pairs were chosen so that they were, as far as possible, neither visually nor phonologically confusable. Spring's work (1971) on the visual similarity of letters provided guidelines for choice. Further, no letter pairs which could be pronounced were used (e.g., *OB*; *At*).

The letter pairs were printed centrally on white card with black, 84 pt. Helvetica medium letraset and presented in a tachistoscope (Campden Instruments Model 610). The presentation of a letter pair activated a millisecond timer (Forth Instruments Ltd., Model F1 114) and the subject's vocal response stopped the timer by means of a voice key (Campden Instruments, Model 340).

FIGURE 1

TEST LETTER PAIR STIMULI AND CONDITIONS

Visual Code Pairs			
phys same	phys. same rev.	phys. diff	phys. diff. rev.
ss	bb	sr	bd
tt	dd	rz	db
nn	pp	ns	pq
kk	qq	mv	qp
Name Code Pairs			
name same	name same rev.	name diff	name diff. rev.
Rr	Dd	Rs	Db
Hh	Bb	Hs	Bd
Gg	Pp	Gw	Qp
Tt	Qq	Tm	Pq

If it was found that the DYS group did not differ from the CAC group on a particular function, it would be inferred that no deficiency existed for dyslexic children in that function. If they differed from the CAC group but not from the RAC group, then quite possibly their behaviour on that particular function could be interpreted as indicative of a deprivation of exposure to written language usually experienced by a child of that age. If the DYS group were found to differ, either qualitatively or quantitatively, from both the CAC and the RAC groups on a particular function, then it could be tentatively concluded that the area of dysfunction made a causal contribution to dyslexia (Seymour and Porpodas, 1980).

Procedure

Testing occurred in a quiet classroom. It was explained to the children that they would be seeing pairs of letters printed on white card. At first, these would be held in front of them, but later they would see them through a tachistoscope. If both the letters in the pair had the same name, the children were instructed to respond 'same', if they had a different name they were instructed to respond 'different'. They were also asked to give their answer as soon as they knew it. The children were then given 12 practice trials using flash cards (not test stimuli) and corrected if they made an error, e.g., "No, that's not right because both letters have the same name".

Twelve further trials, using these same flash cards, were then given using the tachistoscope. Each trial was preceded by a ready signal and feedback on error was given as before. The test cards were then shown using the tachistoscope. No feedback was given but the children were encouraged from time to time by praising their behaviour. A rest period of 5 minutes, midway through the cards, was provided when the children were allowed to move around the classroom, chat and offered chocolates. The test card stimuli were presented in quasi-random order, with no more than 3 'same' or 3 'different' responses occurring consecutively (as Ellis, 1981). The total number of 'same' and 'different' responses required by the task were the same. Both speed (response latencies) and error data were collected.

A baseline reaction time measure was collected for each subject under the same conditions as, and prior to, the test stimuli. A two-choice reaction time task was chosen and the child was instructed to say 'yes' after the onset of a pattern and 'no' if the pattern was absent but a black cross appeared. Practice was given as before, using flash cards, followed by five practice trials in the tachistoscope and finally the four test trials.

RESULTS

Analysis of basic reaction times

No errors were reported for any of the groups. The baseline reaction times were analysed by ANOVA. The 'group' factor was significant (F = 32·38, df = 2, P < 0·001). Inter group t-tests showed that while differences between the DYS and CAC groups were non-significant the RAC group were slower than both the DYS group (t = 7·23, df = 78, P < 0·001) and the CAC group (t = 5·143, df = 78, P < 0·001). The means and SDs are shown in Table 2.

TABLE 2

BASELINE REACTION TIME MEANS (IN SECONDS) AND STANDARD DEVIATIONS

	Dyslexic	CAC	RAC
Mean	0 6023	0·6291	0·8069
SD	0·1461	0·1279	0·1607

Error data

TABLE 3

NUMBER OF ERRORS MADE BY EACH GROUP ON EACH CONDITION

	Visual Code Pairs					Name Code Pairs				
Cond	phys same	phys same rev	phys diff	phys diff rev	Total Visual Error	name same	name same rev	name diff	name diff rev	Total Name Error
DYS	0	1	3	9	13/160	9	16	7	14	46/160
CAC	2	1	4	3	10/160	2	11	3	11	27/160
RAC	1	2	2	7	12/160	4	12	0	13	29/160

Any significant increase in the number of errors made on processing reversible letter pairs relative to the corresponding non-reversible letter pairs, was taken to be a measure of 'reversibility confusion'.

Kruskal-Wallis one-way ANOVA's (2 conditions; reversible/non-reversible) for each group, type and response, showed that, in terms of error rate, there was no reversibility confusion in the visual code for any of the experimental groups. For mixed case stimuli, all groups made markedly more errors when these included reversible letters (Table 3), though reversibility confusion, as measured here, was only significant for the CAC and RAC groups. This possibly reflects the fact that the DYS group made significantly more errors than either of the control groups on all mixed case stimuli (Kruskal-Wallis one-way ANOVA, 3 groups with 10 subjects in each, $\chi=8\cdot125$, $P<0\cdot01$). Further, an impulsivity effect was noticed for the CAC children during the experimental procedure; a trade-off of speed for accuracy.

TABLE 4

SIGNIFICANCE OF ERROR INCREASE ON ADDITION OF REVERSIBILITY

		DYS	CAC	RAC
Visual Code (Lower Case)	phys same/ phys same rev.	—	—	—
	phys diff/ phys diff rev.	—	—	—
Name Code (Mixed Case)	name same/ name same rev.	—	**	*
	name diff/ name diff rev	—	*	***

* $P<0\cdot05$; ** $P<0\cdot01$; *** $P<0\cdot001$
(Kruskal-Wallis one-way ANOVA)

Response Time data

RTs were analysed as a 4-way ANOVA with subjects nested within groups: 3 groups (DYS, CAC, RAC) × 2 types (mixed case, lower case) × 2 response (same, different) × 2 reversibility (reversible, non-reversible). Results are shown in Table 6.

TABLE 5

MEAN LATENCIES OF RESPONSE (IN SECONDS) FOR THE THREE EXPERIMENTAL GROUPS

	Visual Code pairs				Name Code pairs			
Cond	phys same	phys same rev	phys diff	phys diff rev	name same	name same rev	name diff	name diff rev
DYS	0·9748	1·0286	1·0520	1·2377	1·1230	1·4471	1·1923	1·4786
CAC	0·8605	0 8895	0 9503	1·1472	0·9950	1·2000	0·9882	1·0990
RAC	1·1970	1·5129	1·5478	1·5916	1·5290	1·9472	1·6317	1·8972

TABLE 6

ANOVA ANALYSIS RESULTS

Source of variation	df	MS	F value	Sig. of F
Between Groups	2	29·350	146·822	0·001
Between Type	1	10 869	54·371	0·001
Between Response	1	2·077	10·389	0·001
Between Reversib.	1	8·782	43·933	0·001
Group × Type	2	0·723	3·615	0·027
Group × Response	2	0·064	0·321	0 725
Group × Reversib.	2	0·206	1·033	0·357
Type × Response	1	1·830	9·154	0·003
Type × Reversib.	1	0 711	3·555	0 060
Response × Reversib.	1	0·036	0·181	0 670

Within the present framework, 'between reversibility' and the 'type × reversibility' interaction were the most interesting sources of variance. However, before focusing the discussion on these two effects, the remaining effects and interactions will be briefly explained. First, significance of the 'between groups' factor indicates that the three groups of children are responding differentially to the stimuli. As Figure 2 shows, the CAC children responded the fastest, followed by the DYS children and lastly the RAC children. Second, both 'between type' and 'group × type' interaction were significant. This expected result reflects the Posner effect, i.e., all three groups respond faster to same case, visual code matches than to mixed case, name code ones. Third, 'between response' and 'type × response' interaction were significant effects, reflecting the greater amount of processing required by 'different' responses compared to 'same' responses, an effect particularly marked for visual code matches (Figure 2).

As 'reversibility' was significant as a main effect, it can be concluded that, overall, stimuli containing reversible letters take longer to process than those which do not. Further, as the interaction 'type × reversibility' approached significance at the 5 per cent level, there would seem to be a tendency for the lower and mixed case matches to be differentially affected by the presence of reversible letters though, as a glance at Table 5 shows, only the RAC group is affected by reversibility confusion in the processing of physically identical pairs.

T-tests were therefore computed to discover the extent and locus of any reversibility confusion (i.e., any significant increase in processing time for reversible letter pairs relative to the corresponding non-reversible letter pairs), for each group, type and response. These are shown in Table 7.

FIGURE 2

MEAN RTS (IN MSEC) V CONDITION FOR VISUAL AND NAME CODE CURVES

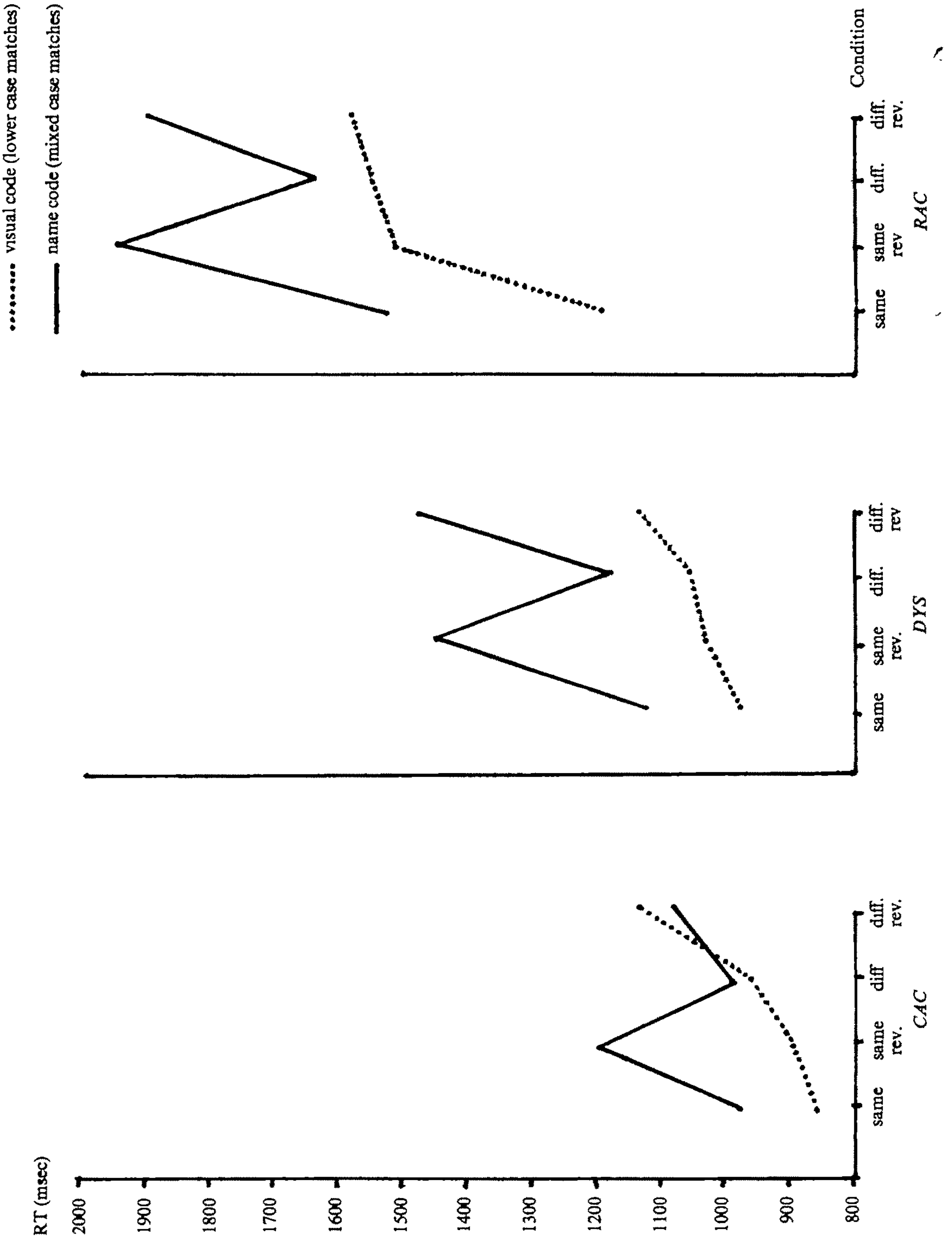


TABLE 7

LEVELS OF SIGNIFICANCE OF REVERSIBILITY CONFUSION FOR EACH GROUP, TYPE AND RESPONSE

		DYS	CAC	RAC
Visual Code	phys same/ phys same rev.	—	—	**
	phys diff/ phys diff rev.	**	**	—
Name Code	name same/ name same rev.	**	*	*
	name diff/ name diff rev.	**	—	—

* P<0·05; ** P<0·01

DISCUSSION

Both speed and error data indicate that additional processing is involved in the 'different' response to physically different stimuli and so it would seem that the time to respond 'same' to physically identical stimuli gives the purest measure of the act of perception. As Table 7 and Figure 2 show, the processing of reversible letters is very similar for both DYS and CAC children for visual code matches. Dyslexic children literally 'perceive' reversible letters as well as, or almost as well as, they perceive non-reversible letters, and as well as normal readers of the same chronological age perceive them. Dyslexic children function well in the visual code and this code does not appear to be the locus of their reversible letter confusions.

Whilst all three groups show some measure of reversibility confusion when processing mixed case stimuli, the dyslexic children show more consistent confusion, in terms of RT, than either of the control groups (Table 7). Whether Posner tasks involve lexical access (i.e., the activation of letter logogens or pattern recognisers for letters), the application of grapheme-phoneme correspondence rules, or the matching of 'abstract letter identities' (Coltheart, 1981), is undecided. But reversibility confusions do seem to be mainly occurring somewhere during the translation from the visual code to either the name code or its referrents, and to be more pernicious in dyslexic readers.

Two further observations are worthy of note. First, only dyslexic children show any reversibility confusion during processing of name different reversible letter pairs. Second, the CAC children process name different reversible letter stimuli (eg., *Bd*) more quickly than they process physically different reversible letter stimuli (eg., *bd*), while for both the DYS and RAC groups all visual code matches are quicker than the corresponding name code matches. Ellis (1981) also found shorter mean reaction times for his 'phonologically confusable different' condition than for his 'physically confusable different' condition, for normal readers but not for dyslexic children (though phonological and reversible confusion were confounded here). Remembering that the reversible letters are also phonologically similar, one possible explanation of both these observations involves a facilitative effect produced by phonological similarity of the stimuli, present in normal readers but not in dyslexic children. Bigsby (1984) found that phonologically similar mixed case letter pairs (e.g., *Cv*) were processed significantly faster than other mixed case letter pairs by CAC children (P < 0·05) though not by dyslexic children. Therefore, research programmes designed to investigate the effects of phonological confusability on visual information processing should take care not to confound phonologically confusable letters with reversible ones.

This study concludes, in agreement with Vellutino (1977), that reversible letter confusions are largely errors of labelling and only minimally those of perception, at least for children of 9·5-11·5 years, though perhaps the balance is different for younger readers. If this is so, then remedial programmes aimed at improving the visual perception of dyslexic children would seem to be of little value and, in fact, such programmes have been shown to be unproductive (ZuWallack, 1977). Structured procedures which stress letter to sound mappings and are supported by extra input routes (e.g., kinaesthetic ones) might be profitable (Miles and Ellis, 1981). New developments in voice synthesiser chips provide an exciting range of possibilities here.

Acknowledgments. — The author would like to thank Mr. Bob Thompson, Senior Educational Psychologist, County Hall, Norwich, the staff and children of Norwich schools for their cooperation, and Mr. Steve Mosely for help with the analysis.

REFERENCES

Bigsby, P (1984). Dyslexia defined as a functional deficiency in visual information processing, at the name code level. MA dissertation. University of East Anglia, Norwich.

Coltheart, M. (1981). Disorders of reading and their implications for models of normal reading. *Visible Language*, 15, 245-286.

Critchley, M. (1970). *The Dyslexic Child* Springfield, Ill.. Charles C. Thomas.

Denckla, M B., and Rudel, R. G (1976). Naming of object-drawings by dyslexic and other learning disabled children. *Brain and Language*, 3, 1-15

Ellis, N. (1981). Visual and name coding in dyslexic children. *Psychol. Res.*, 43, 201-218.

Ellis, N , and Miles, T. (1977) Dyslexia as a limitation in the ability to process information *Orton Society Bull.*, 27, 72-81.

Ellis, N , and Miles, T. (1981). A lexical encoding deficiency I· Experimental evidence. In Pavlidis, G. Th., and Miles, T. R. (Eds), *Dyslexia Research and its Implications to Education*. Chichester· Wiley.

Haber, R. N. (1969). *Information Processing Approaches to Visual Perception*. New York: Holt Rinehart

Hawkins, H. L. (1969). Parallel processing in complex visual discrimination. *Percept Psychophysics*, 5, 56-64.

Jorm, A. F. (1979). The cognitive and neurological basis of developmental dyslexia. a theoretical framework and review *Cognition*, 7, 19-33

Katz, L., and Wicklund, D. A. (1971) Word scanning rate for good and poor readers *J educ. Psychol.*, 62, 138-140.

Legein, Ch. P., and Bouma, H (1981). Visual recognition experiments in dyslexia. In Pavlidis, G. Th., and Miles, T R. (Eds.), *Dyslexia Research and its Implications to Education*. Chichester: Wiley.

Liberman, I , Shankweiler, D , Orlando, C., Harris, K , and Berti, F. B. (1971). Letter confusions and reversals of sequence in the beginning reader· implications for Orton's theory of development dyslexia. *Cortex*, 7, 127-142.

Lyle, J. G., and Goyen, J. D. (1975). Effect of speed of exposure and difficulty of discrimination on visual recognition of retarded readers. *J. abnorm. Psychol.*, 84, 673-676.

Miles, T. R., and Ellis, N C. (1981). A lexical encoding deficiency II: clinical observations. In Pavlidis, G. Th., and Miles, T. R. (Eds.), *Dyslexia Research and its Implications to Education*. Chichester· Wiley.

Orton, S T (1925). Word-blindness in school children. *Arch. Neurol. Psychiat.*, 14, 581-615

Posner, M. I , and Mitchell, R. F. (1967). Chronometric analysis of classification. *Psychol. Rev* , 74, 392-409

Seymour, P. H. K., and Porpodas, C. D. (1980) Lexical and non-lexical processing in developmental dyslexia. In Frith, U. (Ed.), *Cognitive Processes in Spelling*. London. Academic Press.

Spring, C (1971). Perceptual speed in poor readers. *J. educ. Psychol.*, 62, 492-500

Stanley, G., and Hall, R (1973). Short-term visual information processing in dyslexics. *Child Dev.*, 44, 841-844.

Vellutino, F. R (1977). Alternative conceptualizations of dyslexia: evidence in support of a verbal deficit hypothesis. *Harv. Educ. Rev* , 47, 334-354.

Vellutino, F. R., Steger, J. A., and Kandel, G. (1972). Reading disability· an investigation of the perceptual deficit hypothesis. *Cortex*, 8, 106-118

ZuWallack, R J (1977). The effects of visual sequential memory training upon visual sequential memory and reading achievement in first grade *Dissert. Abstr. Int.*, Order no. 77-26, 586, 334 pages.

(*Manuscript received on* 11*th February,* 1985)

Br. J. educ. Psychol., 55, 273-279, 1985

VERBAL/SPATIAL TASK EXPECTATIONS IN WRITTEN INSTRUCTION

By GOLDIE S. KRAFT, GREGORY J. NOVIE, AND RAYMOND W. KULHAVY
(*Arizona State University*)

SUMMARY. In two experiments, high school students studied a map–prose combination under instructions to expect either a spatial/locational or verbal/event post-test over the target material. Control groups received no expectation instructions. In general, test expectation instructions significantly influenced the kinds of information students were able to recall on a spatially or verbally cued post-test. In the first study, people who expected a verbal test recalled significantly fewer location items, whereas people expecting a spatial test produced superior location recall. In the second study, students in both verbal and spatial expectation groups outperformed the control condition on specifically cued items. These data suggest that the type of instructional test expectation a student is exposed to can have a major effect on the way in which instruction will be processed and studied

INTRODUCTION

EXACTLY what people learn and remember when they read an instructional unit is strongly influenced by what they expect to do with the information after reading. Variables such as test-like events, study objectives, and task-specific instructions, have all been used in an attempt to determine how different kinds of task expectations work to affect instructional performance (Duell, 1974; Swenson and Kulhavy, 1974; Kulhavy and Swenson, 1975). One area of emphasis in research on expectation has involved telling students about the form of the test they can expect to receive following study, and then providing them with alternate testing formats in order to assess differences in response performance. For example, Sanjivamurthy and Kumar (1983) gave college learners in an algebra class either a recall or recognition test for six successive weeks, and then changed the mode of testing on the next measure. In general, correct performance suffered when the test mode was changed from recognition to recall; however, students expecting a recall test did about the same on both types of test. Such data indicate that test mode expectation can have a significant effect on how the learner performs following instruction.

In the experiments that follow, we were interested primarily in how task expectations influence test responses which require different types of cognitive activity. Here, our focus was on task demands which emphasised both verbal and spatial representations of the instructional materials, since there are those who have argued that test expectation draws on different cognitive codes in order to accomplish its effect (Tversky, 1973). In our studies, we were able to combine descriptive text passages with related spatial stimuli in the form of a simple reference map. Since the text material is processed in a primarily verbal/linguistic fashion (Anderson, 1978), and evidence is firm that map units are processed spatially in at least a quasi-pictorial manner (Kulhavy *et al.*, 1983), we were willing to predict that student expectations should act to control the way that lesson material is processed into memory. In other words, providing people with a verbal or spatial test expectation should cause them to modify their study strategies so that they process the target instruction in accord with the expectation, and perform better on tests which require a response congruent with what they expected — in this case, either verbal or visual test items.

In Experiment I we reasoned that students who expected a spatial-base test would perform better on a measure which required associations with the reference map elements, and that learners who received verbal test instructions should produce higher scores on a verbally loaded measure. Experiment II was a replication and extension of the first study, except that specific locational referents were substituted for the more general directions in the previous study.

EXPERIMENT I

METHOD

Design and subjects

Instructions telling students what type of post-test to expect were varied between groups to form three separate conditions, with both the type of test and the type of item varied within subjects. The final design was a 3 × test expectation (spatial *vs.* verbal *vs.* control) by 2 × type of test (map *vs.* no map) by 2 × type of item (location *vs.* event) factorial with repeated measures on the last two variables.

The subjects who participated in the first study were 72 ninth-grade students attending high school in Phoenix, Arizona. Twenty-four subjects were randomly assigned to each of the three test expectation conditions.

Materials

Experimental map. The map used in this study was a black-line drawing of the South Island of New Zealand. The map outline was drawn true to scale, reproduced on 8½ × 11 inch (21·25 × 27·50 cm) white paper, and contained in their geographically correct location, 12 labelled towns, one labelled river, two unlabelled lakes and mountain ranges, and one small compass rose positioned in the lower right corner of the map. The 12 towns chosen for their ease of pronunciation and distribution over the map were: Christchurch, Dunedin, Fairlie, Gore, Greymouth, Hermitage, Invercargill, Nelson, Picton, Queenstown, Springs Junction, and Westport.

Text passage. The text used in the experiment was a 600-word narrative written by the authors, and titled, "Beautiful, Historic New Zealand". The passage was written at about the sixth-grade reading level, and was similar in structure to a tour guide presentation. Each of the 12 towns was described according to its referent location on the island map, and each town was associated in the text with one specific historical event which occurred at that spot. The text was printed on 8½ × 11 inch (21·25 cm × 27·50 cm) white paper for use in the study.

Test instruments. Each of the two post-tests consisted of a 12-page booklet, with every page containing both a location and a historical event item for one of the 12 towns mentioned in the passage. The spatial version of the test presented items which consisted of the original map without labels, plus a location and an event item which appeared at the top of the page. All items of the same type had a common stem structure, with only the name of the target town changed from item to item. For example, questions for the town of Gore were:

> Locate the town of GORE on the map by writing its name next to the correct point. (location question)
> Tell what event took place at GORE. (event question)

The written version of the test followed a similar format, except that the items were presented without a map below them. In the case of the written test, the location items were changed to read:

> Write the location of GORE as told in the story. (location question)

The structure of the event items remained the same for both types of tests. The 12 pages for each type of test were stapled into a booklet with a blank green cover sheet, and the order of items in each booklet was separately randomised for each subject.

Procedure

Students participated in small groups with each group completing their task in a separate room to ensure that they would hear only the instructions for their condition. When people in a given group were seated, they each received an instruction sheet appropriate to their condition. Students were told to read the instructions silently while the monitor read them aloud to the group, and all content and procedural questions were answered immediately following the reading. Next, all subjects were asked to turn their instruction sheet over and write out a description of their specific experimental task using their own words. This procedure was followed in order to be sure that all participants understood exactly what was required of them in the study. The monitor checked each of the written descriptions for accuracy, and reviewed missing or incomplete information with the subject prior to collecting the instruction sheets.

Students in all groups were told that they would be reading a short story, studying a map of New Zealand, and receiving two tests over the content once they had finished the learning phase of the experiment. Students in the spatial expectation group were told that they would be tested about historical events taking place near the towns on the map, and that they would also be required to locate the towns themselves as a part of their task. Learners in the verbal expectation condition were given the same instructions concerning events, but were also told that they would need to be able to write out the locations of the towns. Finally, learners in the control group were told only to read carefully, and try to remember as much as possible.

After completing the instruction phase, each learner received a copy of the passage and the map and given a total of 15 minutes in which to study them. Following the learning phase, the materials were collected, and the first of the two tests handed out. The tests were presented in counterbalanced order, with one-half of the subjects in each group receiving the map test first, and the other half receiving the no-map test first. Subjects were given 10 minutes to complete the first test, with a spoken signal given to the group at the 9-minute point. The completed first tests were then collected, and the second test distributed. The time limit and signalling for the second test were identical to the first measure.

RESULTS

The tests were scored according to a predetermined set of criteria, with one point awarded for each correct response, and minor errors in spelling or grammar counted as correct. All statistical tests made on the data were evaluated at the 0·05 level of significance.

An analysis of variance of correctly recalled items (3 test expectations × 2 test types × 2 item types) yielded significance for the test type, $F(1, 69)=194{\cdot}81$, $MS_e=3{\cdot}69$, and item type, $F(1, 69)=4{\cdot}11$, $MS_e=8{\cdot}28$ main effects, and for the test expectation × item type, $F(2, 69)=3{\cdot}61$, MS_e 8·19, and test type × item type, $F(1, 69)=178{\cdot}76$, $MS_e=3{\cdot}00$ interactions. The means and standard deviations contributing to this analysis are shown in Table 1. In general, performance on the map-test was significantly higher than on the no-map test, and correct recall for location items was better than for event items. The test type × item type interaction simply showed that location items were better remembered than event items on the map test, but recall was higher for event items on the no-map measure. The graphic form of the test expectation × item type interaction is diplayed in Figure 1. Tests of simple effects performed on the means contributing to this interaction showed that both the control and the spatial expectation groups recalled significantly more correct location items than the verbal expectation subjects. On event items, both the verbal expectation and the control groups significantly outperformed subjects in the spatial expectation condition.

TABLE 1

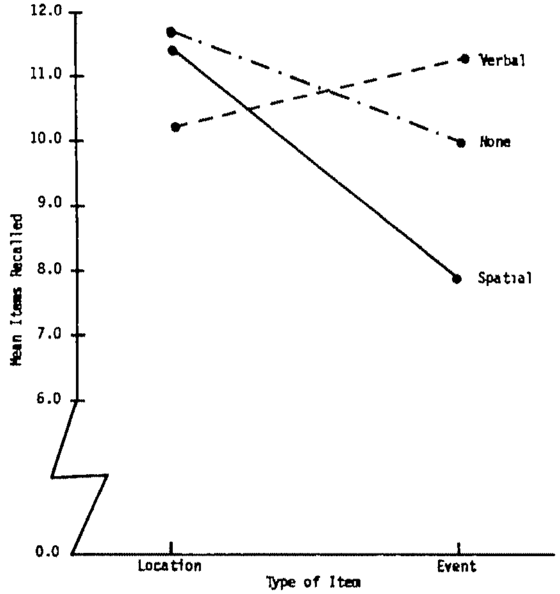
MEAN SCORES AND STANDARD DEVIATIONS FOR CORRECT RECALL: EXPERIMENT I

Expectation Instruction	Map Test				No-Map Test			
	Location Item		Event Item		Location Item		Event Item	
	M	SD	M	SD	M	SD	M	SD
Spatial	9·21	2·78	4·17	3·93	2·21	2·08	3·18	3·67
Verbal	7·46	4·78	5·75	4·16	2·83	2·91	5·46	4·02
None	8·83	3·41	5·33	4·12	2·79	2·64	4·79	3·62

In general, these results show that the type of preinstructional expectation received by a learner has a pronounced effect on how that learner processes the lesson material, and on what he or she is able to remember from study. When test items required students to either write out or identify a specific location, people who expected a verbal test over events recalled significantly fewer location items than either the spatial expectation or control groups. Likewise, when students expected to receive a test which included the spatial elements of the map, there was a dramatic decrease in their performance on event items relative to both the verbal expectation and the control groups — in spite of the fact that event items were identical on all test forms. These data clearly show that the task schema developed by a student prior to instruction can have a powerful influence in determining exactly what content they encode, and how they go about remembering it for a later test.

FIGURE 1

INTERACTION BETWEEN EXPECTATION INSTRUCTIONS AND ITEM TYPE: EXPERIMENT I



EXPERIMENT II

Experiment II was designed as an extension and replication of the first study. We were concerned that the expectation instructions in Experiment I might have had such a strong effect simply because they mainly referenced only one or the other of the spatial/verbal instructional stimuli. Hence, in the second study, the instructions emphasised either locations or events, but stressed that the learner should attend to both the text *and* map in order to obtain the correct target information. Our prediction here was that the location or event expectation would still act to control what type of information a given learner would be most likely to recall, without directing attention specifically to the map or text in the fashion of Experiment I. In order to make a clearer test of our prediction, we also changed the structure of the test items, so that both event and spatial descriptors became item cues, and the 12 town names served as item outcomes for both spatial and verbal response requirements. Finally, in an attempt to eliminate both item prompts and interference, each test item was presented on a separate page, instead of in the pair format used in Experiment I.

METHOD

Design and subjects

Test expectation instructions were again varied to form three between-subjects groups, with both the type of test and the type of item cue treated as within-subjects factors. The structure of the design was thus a 3 × test expectation (location *vs.* event *vs.* control) by 2 × type of test (map *vs.* no map) by 2 × type of item cue (location *vs.* event) factorial with repeated measures on the last two variables.

The subjects were 69 high school students from Phoenix, Arizona, none of whom had participated in the previous study. Twenty-three subjects were randomly assigned to each of the test expectation conditions.

Materials

Map and text passage. Both the stimulus map and the text passage were identical in all respects to those used in Experiment I.

Test instruments. Students received two post-tests, each of which consisted of a 24-page booklet, with one test item on each page. The map version of the test presented each item at the top of the page, with a copy of the original map directly below it without labels. Items having location cues presented a locational description of a town taken directly from the stimulus passage, and students were required to write the name of the town in its correct location on the map. Event-cued items contained the critical information about the historical event occurring in one of the towns, and students were again directed to write the name of the town in its correct position on the map. The no-map version of the test was presented in an identical manner, except that no map appeared on the test pages, and the students were required to write the name of the appropriate town on the blank page below the question. As before, all items possessed a common stem structure, and the order of items in each booklet was separately randomised for each subject.

Procedure

The second study was conducted in the same fashion as Experiment I, except for the content of the instructions given to the different test expectation groups.

People receiving a location test expectation were told that they would be tested on their ability to remember the location of specific towns mentioned in the story and shown on the map. Students in the event expectation group were told that their test would require them to remember information about the historical events mentioned in the text. Instructions for both expectation groups emphasised that they should study both the map and the text in order to be able to respond correctly to the test items.

RESULTS

As before, one point was awarded for each correct test response, and minor structural errors were discounted during scoring. All statistical tests were evaluated at the 0·5 level of significance.

The 3 (test expectation) × 2 (type of test) × 2 (type of item cue) analysis of variance reached significance for both the type of test, $F(1, 66) = 83{\cdot}60$, $MS_e = 6{\cdot}16$, and type of item cue, $F(1, 66) = 67{\cdot}68$, $MS_e = 2{\cdot}12$ main effects, and for the test expectation × type

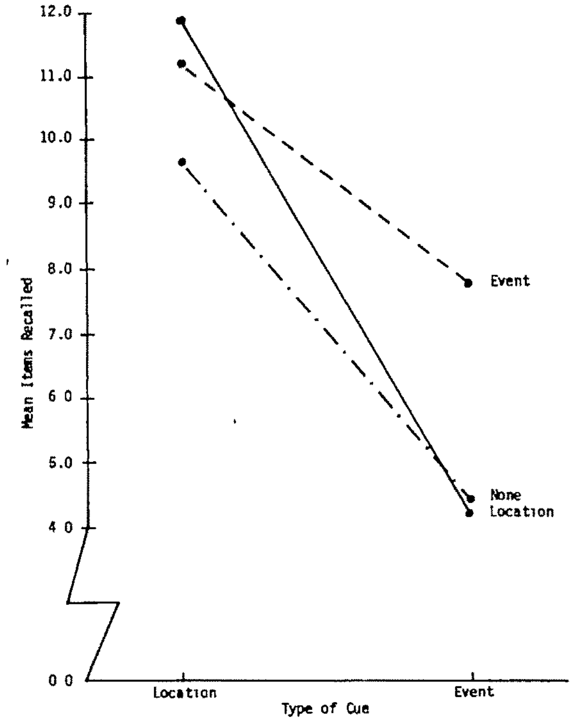
TABLE 2

Mean Scores and Standard Deviations for Correct Recall: Experiment II

	Map Test				No-Map Test			
	Location Cue		Event Cue		Location Cue		Event Cue	
Expectation Instruction	M	SD	M	SD	M	SD	M	SD
Location	7·35	3·97	4·48	4·38	1·91	2·86	2·22	2·81
Event	7·09	3·13	4·09	3·04	3·43	3·41	4·35	3·47
None	6·78	4·08	2·83	3·01	2·17	2·12	2·13	1·98

FIGURE 2

Interaction Between Expectation Instructions and Type of Recall Cue: Experiment II



of test, $F(2, 66) = 4{\cdot}34$, $MS_e = 6{\cdot}16$, and type of test × type of item cue, $F(1, 66) = 177{\cdot}73$, $MS_e = 1{\cdot}97$ interactions. The means and standard deviations for this analysis are displayed in Table 2.

As in Experiment I, correct recall for items on the map test were significantly higher than on the no-map test, and performance on location cued items was better than on event cued items. The significant type of test × type of item cue interaction was due to the fact that location cued items were more often recalled on the map test than event cued items. The form of the test expectation × type of item cue interaction is shown in Figure 2. Simple effect analysis on the means contributing to the interaction showed that both the location and event expectation groups significantly outperformed the controls on items which used a location cue, and then the event expectation subjects recalled significantly more event-cued items than either of the other two groups.

DISCUSSION

The variation in recall patterns across the two studies is probably due to the change in the type of test activity required of the students. In spite of such differences, both of the experiments clearly favour our initial assumption, that instructions to expect a specific test item domain influence the way in which text information is processed and remembered. Students who expected a spatial/locational test recalled significantly more perceptually related items than their verbal/event counterparts in Experiment I, and significantly more such items than the controls in Experiment II. Similarly, students receiving a verbal/event expectation reliably outperformed the spatial/locational groups in both studies, and the non-instructed controls in Experiment II.

Our results have some useful implications for the classroom instruction and learning process. From the data from these studies, we are willing to assume that how students go about the business of studying lesson material is, at least, partially a function of the form and domain of information over which they expect to be tested. Hence, in addition to the usual recommendation that tests should reflect critical elements of instruction, we can add the proviso that learners should be carefully told about exactly what kind of cognitive process will be represented during the measurement phase of the lesson. Such specific directions will have the dual advantage of keying in appropriate processing tactics, and of increasing the probability that a test will enable to differentiate between those students who have acquired important content, and those who have not.

Reprint requests should be sent to. Dr. Raymond W. Kulhavy, Department of Educational Psychology, Arizona State University, Tempe, Arizona 85287, U.S.A.

REFERENCES

ANDERSON, J. R. (1978). Arguments concerning representations for mental imagery. *Psychol. Rev.*, 85, 249-277.

DUELL, O. K. (1974). Effect of type of objective, level of test questions, and the judged importance of test material on post-test performance. *J. educ. Psychol.*, 66, 225-232.

KULHAVY, R. W., SCHWARTZ, H. H , and SHAHA, S. H. (1983). Spatial representation of maps. *Am. J. Psychol.*, 96, 337-351.

KULHAVY, R. W., and SWENSON, I. (1975). Imagery instructions and the comprehension of text. *Br. J. educ. Psychol.*, 45, 47-51.

SANJIVAMURTHY, P. T., and KUMAR, V. K. (1983). Test mode anticipation and performance: A classroom experiment. *Contemp. educ. Psychol.*, 8, 355-365.

SWENSON, I., and KULHAVY, R. W. (1974). Adjunct questions and the comprehension of prose by children. *J. educ. Psychol.*, 66, 212-215.

TVERSKY, B. (1973). Encoding processes in recognition and recall. *Cogn. Psychol.*, 5, 275-287.

Br. J. educ. Psychol., 55, 280-287, 1985

THE PREDICTIVE VALIDITY OF BANNATYNE'S WISC CATEGORIES FOR LATER READING ACHIEVEMENT

By MICHAEL F. McKAY, MARIE D. NEALE AND G. BRIAN THOMPSON
(*Monash University, Victoria, Australia*)

SUMMARY. This study was concerned with the Preparatory Grade prediction of later reading performance using recategorised WISC results suggested by Bannatyne (1974). Using the three factors Spatial, Conceptual and Sequential as the dependent variables, and group membership based on subsequent longitudinal reading performance as the independent variables, a multivariate analysis of variance (MANOVA) resulted in a significant multivariate F for a sample of 179 children ($P < 0.0001$). Step-down tests and post-hoc comparisons indicated that the Sequential factor accounted for most of between-group variance and was the only factor to discriminate between the 'failing readers' and average and above average readers. None of the Bannatyne-WISC categories was able to discriminate between children who were 'slow starters' in beginning reading and children who continued to fail at reading throughout the primary school.

INTRODUCTION

The Wechsler Intelligence Scale for children, WISC (Wechsler, 1949) has been the object of a great deal of empirical investigation since its origin (Glasser and Zimmerman, 1967; Zimmerman and Woo-Sam, 1972; Sattler, 1974). A large part of this research has concerned itself with the psychoeducational utility of using the WISC in the diagnostic process (Dockrell, 1960; Huelsman, 1970; Egeland *et al.*, 1970; Ackerman *et al.*, 1971; Levine and Fuller, 1972; Fuller and Friedrich, 1974). A number of studies have attempted to identify characteristic subtest patterns on the WISC for poor readers compared to average or above average readers. However, the question of whether characteristic subtest patterns on the WISC for poor readers, after several years of schooling, are an antecedent or consequence of reading disability has not been addressed. It seems important to know if WISC patterns of performance for poor readers are simply *associated* with factors known to contribute to failure or are indicative of underlying deficits which affect the individual's ability to learn to read. If the latter, it would be also useful to know how early in development such deficits are revealed by the WISC and what are the longer term prospects of such difficulties.

One approach which could help clarify this difficulty would be to relate subtest scores on the WISC to a factor structure which has some psychological meaning, as far as the reading process is concerned, rather than subtest performance *per se*. Such an approach has been put forward by Bannantyne (1968, 1974) in his scheme for the recategorisation of WISC scores. He suggested that a tripartite organisation of the WISC scores into categories called Spatial, Conceptual and Sequential, possessed more psychological meaning and diagnostic utility than conventional subtest analyses. According to Bannatyne, subtests in the Spatial category (Block Design, Object Assembly and Picture Completion) require the ability to manipulate objects directly or symbolically in multi-dimensional space. Subtests in the Conceptual category (Vocabulary, Similarities and Comprehension) involve utilisation of concepts and abstract reasoning. Subtests in the Sequential category (initially Digit Span, Coding and Picture Arrangement, but later reorganised to include Arithmetic and omit Picture Arrangement) require the ability to retain sequences of auditory and visual stimuli in short-term memory storage. A number of studies have substantiated this factor structure of the WISC including Rugel (1974b), Kaufman (1975) and McKay (1977) among others.

Further, Bannatyne (1971) has reported that 'dyslexic' readers received their highest scores in the Spatial category, intermediate scores in the Conceptual category and their lowest scores in the Sequential category. Rugel (1974a) found this same pattern with a broader spectrum of disabled readers. He reanalysed the data from over 20 studies involving the use of the WISC with disabled readers and found that, as a pooled group, the children did indeed manifest recategorised scores with the ranking Spatial > Conceptual > Sequential. Smith *et al.*, (1977), in assessing the utility of Bannatyne's recategorisation of WISC-R subtest scaled scores with a population of school-verified learning disabled subjects, found further support for such a recategorisation. The learning disabled subjects were characterised by the same pattern of abilities that Bannatyne (1968, 1971) found for 'dyslexics' and that Rugel (1974a) reported for disabled readers in general, i.e., Spatial > Conceptual > Sequential. Children with Full Scale IQ of 75 or below failed to show this pattern.

However, although it could be argued that the Bannatyne recategorisation may have more psychological meaning than conventional diagnostic categories for the WISC — in that the factors identified have been related to reading failure — the question of causation has not been addressed in previous research. The age groups of the subjects in the studies cited have involved children already several years into primary school whose test performance may be contaminated by school related variables which might be convarying with poor WISC performance. Schiff *et al.*, (1981) and Kaufman (1981) have argued that Bannatyne's Sequential factor appears to be synonymous with a factor typically described as 'Freedom from Distractibility' and concerned with attention to the task, concentration or anxiety. Such a set of behaviours could be constitutional (hence be evident before school) or could arise from frustrating schooling experiences and develop only after schooling has commenced. There are ample studies in the literature characterising children with learning difficulties as having an inability to focus attention and adopt appropriate perceptual strategies (e.g., Dykman *et al.*, 1970, 1971; Hynd *et al.*, 1978, 1979; Obrzut *et al.*, 1981). The question as to whether the pattern of WISC performance Spatial > Conceptual > Sequential is able to discriminate effectively, in the long term, between those children whose retardation in reading performance is due to exogenous factors and those whose retardation is due to some critical underlying ability therefore bears investigating.

The primary purpose of this study was to investigate (*a*) if children who continue to fail at reading in primary school (despite overall average intelligence) display a deficit in WISC results characterised by Bannatyne's Sequential factor and, (*b*) if they display the characterised pattern of Spatial > Conceptual > Sequential reported by Bannatyne (1971) and Rugel (1974a).

METHOD

A random sample of 6 per cent of the pre-schools on the register of the Division of Maternal and Child Welfare of the Department of Health, Victoria, was identified for this study. From their enrolments 20 per cent (106 boys and 115 girls) were then randomly selected. The mean age of subjects at time of selection was 4.6 years. The sample was approximately representative of the distribution of SES in the general population, as measured by father's occupation (Congalton, 1969), with 32.9 per cent from professional categories (1,2,3), 49.5 per cent from small business, self-employed categories (4,5,6) and 17.6 per cent from semi-skilled and unskilled categories (6,7). Intellectual ability when measured by the WISC some 18 months after study had commenced indicated an average mean IQ for the sample of 109.17 (SD 12.57), with mean Verbal IQ and mean Performance IQ of 108.4 and 108.10 respectively.

Procedure and data analysis

Data were collected on the sample from kindergarten to Grade 4 (approximate age 4.6 years to 10.6 years), at about 12-monthly intervals over the period 1970-1975. During this time 42 children from the original sample were lost primarily due to moves of school or unavailability of group test scores in one or more years. Therefore, a final sample of 179 were used in the analysis. The data collected included the WISC and measures of reading performance, the Neale Analysis of Reading Ability (Neale, 1958) and the ACER Reading Survey Test, FORM B (ACER, 1972). The first reading test is an individually administered set of graded narratives yielding scores for Accuracy (word recognition), Comprehension and Reading Speed with a categorisation of the errors made. The second reading assessment, (ACER Primary Reading Survey Test, Comprehension, Level B, Form R) is a measure of independent silent reading and was administered by the subjects' classroom teachers according to written instructions provided by the present authors.

Given that repeated measures of reading ability were available over a 4-year period in this study and given that the authors were interested in changes in reading performance over a long period of time (5 years), it seemed plausible to depart from traditional multivariate forms of analysis e.g., multiple regression. One of the criticisms of predictive studies has been that although it can be shown that criterion groups may differ on predictive variables, it becomes more difficult to apply such information when the developmental characteristics of different individuals are being considered. If one can identify the characteristics which distinguish between those early 'at risk' children who ultimately fail from those who do not, then early diagnosis becomes more accurate.

This departure from traditional analysis was achieved by taking standardised reading assessments in the form of stanine rankings, in each of a number of successive years, and plotting patterns of reading performance in terms of relative change from one stanine group to another. The distributions of reading scores for each year were checked in order to establish that they were normal.

It thus became possible to categorise readers into eight discrete groups based on 'change' or 'no change' criteria relative to the position of other members in the sample from years 1-4 inclusive. These groups were a *Low-low* group or "continually failing" group who were in the lowest 3 stanines in all years assessed, N = 24; a *Low-up* group or "slow starters" who were in the lowest 3 stanines in all years assessed, N = 16, a *Medium* group or "average" group (middle stanines 4, 5, 6 all years assessed, N = 64), a *Medium-down* group (initially stanines 4, 5 or 6, but lower in subsequent years, N = 11), a *Medium-up* group (middle stanines first year assessed, higher in subsequent years N = 21), a *High-down* group (top 3 stanines when first assessed, lower rating in subsequent years, N = 16), a *High-high* group "above average" readers (maintained membership in top 3 stanines in all years assessed N = 16) and finally a Variable group whose performance was rather more erratic across the period of study, N = 11. The same basic procedures are outlined more fully elsewhere (McKay, 1977; Neale *et al.*, 1979).

RESULT AND DISCUSSION

The question of differences between the reading groups on the three WISC categories was assessed using a multivariate analysis of variance (MANOVA) model to determine the significance of the overall differences within and between the eight groups described on the predictor variables. Because of the nature of the data concerning the independent variable, i.e., group membership instead of individual reading scores, MANOVA was preferred as the general data analytic technique.

Multivariate analyses were undertaken for boys and girls separately and for the total sample. Overall F-ratios for the three analyses indicated significant differences between the eight reading groups on the three dependent variables Spatial, Conceptual and Sequential for boys (multivariate F value = 2.5, df 21/213, $P<0.001$), for girls (multivariate F value = 3.45, df 21/244, $P<0.0001$) and for the total sample (multivariate F value = 4.51, df 21/485, $P<0.0001$). Univariate F-tests were then conducted to see whether the overall significant differences occurred for *each* dependent variable. Table 1 shows the univariate F-ratios for each variable for the total sample, boys and girls.

TABLE 1

UNIVARIATE AND STEP-DOWN F-RATIOS WITH 'SEQUENTIAL' AS FIRST ORDERED VARIABLE

Variable	Hypothesis mean squares	Univariate F	P less than	Step-down F	P less than	
Total N = 179	df 7/171					
1. Sequential	398·54	12·91	0·0001	12·91	0·0001	
2. Conceptual	248·76	4·47	0·0002	1·59	0·1424	NS
3. Spatial	81·66	2·88	0·0072	0·67	0·6960	NS
Boys N = 84	df 7/76					
1. Sequential	223·46	5·79	0·0001	5·79	0·0001	
2. Conceptual	115·91	1·74	0·1112	1·41	0·2130	NS
3. Spatial	65·46	1·92	0·0782	0·90	0·5075	NS
Girls N = 95	df 7/87					
1. Sequential	197·11	8·09	0·0001	8·09	0·0001	
2. Conceptual	179·64	3·79	0·0013	1·98	0·0676	NS
3. Spatial	56·70	2·51	0·0211	1·16	0·3351	NS

Significant differences occurred for each dependent variable among the eight reading groups in the case of the total sample and for the girls. However, neither the Conceptual variable (univariate F value 1.74, df 7 & 76) nor the Spatial variable (univariate F value 1.92, df 7 & 76) discriminated significantly among the eight reading groups when the boys' scores were considered separately.

The relative effects of each dependent variable in contributing to the between-group's variance was then investigated by means of the step-down test (Roy and Bargmann, 1958). Given that the variable of crucial interest was the Sequential one it was placed first in the step-down test, enabling the effects of the other two dependent variables to be partialled out (see Table 1). When this was carried out there was no remaining residual variance related to the Conceptual and Spatial variables. This applied to the results of the total sample, boys and girls. Only the Sequential variable, therefore, demonstrated an overall effect independent of the other two variables in the set.

Multiple comparisons using Fisher's (1949) modified LSD test were then carried out to investigate all between-group differences on each variable (see Table 2).

Of the 16 comparisons which indicated significant differences, 11 occurred on the Sequential factor.

The Low-low group (the continuously failing readers) was significantly different on this factor from all the other reading groups except the Low-up group (the slow starters). While the continually failing readers differed significantly from the above average readers (the High-high group) on all three WISC factors, significant differences between this group and the average readers (the Medium group) occurred only on the Sequential variable. The observation by Rugel (1974b) that low scores on

TABLE 2

RESULTS OF POST-HOC COMPARISONS FOR THE THREE WISC FACTORS, SPATIAL, CONCEPTUAL, SEQUENTIAL

Contrast	Sequential	Conceptual	Spatial
1. Low-low vs Low-up	NS	NS	NS
2. Low-low vs Medium	*	NS	NS
3. Low-low vs High-down	*	NS	NS
4. Low-low vs High-high	*	*	*
5 Low-low vs Medium-up	*	*	NS
6 Low-low vs Medium-down	*	NS	NS
7 Low-low vs Variable	*	NS	NS
8 Low-up vs Medium	*	NS	NS
9. Low-up vs High-down	*	NS	NS
10. Low-up vs High-high	*	*	NS
11. Low-up vs Medium-up	*	*	NS
12. Low-up vs Medium-down	NS	NS	NS
13. Low-up vs Variable	*	NS	NS
14. Medium vs High-down	NS	NS	NS
15. Medium vs High-high	NS	NS	NS
16 Medium vs Medium-up	NS	NS	NS
17. Medium vs Medium-down	NS	NS	NS
18. Medium vs Variable	NS	NS	NS
19. High-down vs High-high	NS	NS	NS
20. High-down vs Medium-up	NS	NS	NS
21. High-down vs Medium-down	NS	NS	NS
22 High-down vs Variable	NS	NS	NS
23 High-high vs Medium-up	NS	NS	NS
24. High-high vs Medium-down	NS	NS	NS
25. High-high vs Variable	NS	NS	NS
26. Medium-up vs Medium-down	NS	NS	NS
27. Medium-up vs Variable	NS	NS	NS
28. Medium-down vs Variable	NS	NS	NS

* P<0 05

the Sequential category would distinguish continually failing readers from consistently average or good readers was, therefore, supported.

The predictive validity of the Sequential factor would have been strengthened if it had discriminated between the continually failing readers (the Low-low group) and the slow starters (the Low-up group). However, the only group that the slow starters did *not* differ from on the Sequential variable was the continually failing group. Low performance on the Sequential variable in their preparatory year at school was thus a characteristic of failing readers *and* early below average readers who improved in reading performance later in their schooling. It would be convenient to speculate on the possibility that the slow starters (those who subsequently improved in reading) were the group who were able to overcome their distractibility and poor attentional skills early in schooling, and that for this reason they improved. However, as there are no data (in the form of WISC retest scores or teacher ratings, for example) to support this observation, it remains an open question. Obviously low performance on the Sequential factor early in the child's school career (i.e., around 6.0 years of age) is not, in itself, sufficient reason to indicate a specific difficulty in acquiring reading skills. However, if it occurs along with a pattern of high distractiblity, inattentiveness and low concentration in general behaviour then it could be quite significant.

The second question addressed in this study concerned the Spatial > Conceptual > Sequential pattern of disabled readers on the WISC. Although, as an independent predictor of later reading performance, the Sequential factor was shown to have some validity by the findings reported above, it was not able to discriminate between those who were slow starters at learning to read from those who had continuing reading problems. However, was there a *pattern* on the WISC categories for the failing group resembling Rugel's reported pattern which characterised this group from the other readers? Table 3 shows the means and standard deviations for each group on each of the three WISC factors, Spatial, Conceptual and Sequential.

TABLE 3

MEANS AND STANDARD DEVIATIONS FOR EIGHT READING PERFORMANCE GROUPS ON THREE WISC FACTORS, SPATIAL, CONCEPTUAL, SEQUENTIAL (TOTAL N = 179)

Group		Spatial		Conceptual		Sequential	
	N	$\bar{x}$	SD	$\bar{x}$	SD	$\bar{x}$	SD
1 Low-low	(24)	31·33	6·11	28·00	8·67	25·08	5·31
2 Low-up	(16)	31·88	4·73	26·94	7·07	27·69	6·60
3. Medium	(64)	34·38	5·82	33·44	7·12	33·69	5·76
4. Medium-down	(11)	31·55	4·65	31·00	5·81	32·36	6·52
5 Medium-up	(21)	34·71	4·72	36·95	7·19	34·19	4·35
6. High-down	(16)	34·25	4·25	33·44	8·35	35·25	5·90
7 High-high	(16)	37·88	5·10	36·44	8·25	39·25	4·11
8. Variable	(11)	33·27	4·34	32·09	6·43	35·64	5·55

An inspection of Table 3 indicated that the pattern Spatial > Conceptual > Sequential was indeed manifest for the failing reader group (the Low-low group) and of the eight groups identified in this study was unique to this group, at least superficially. However, a single classification repeated measures analysis of variance for the Low-low group on the three sets of results Spatial, Conceptual and Sequential resulted in a non-significant F value (7.41 df 2/23), thus failing to support the notion that a significant difference in the "pattern of means" for these three factors.

Summary

The findings of this study, therefore, add some support to the claim that poor performance on the Sequential category of the WISC, characterises failing readers. The further finding, that poor performance on the Sequential category of the WISC as early as 6.0 years is *predictive* of later reading failure, should also be added. This latter finding, however, needs to be modified in view of the fact that poor performance on the Sequential category, early in their school career, is also characteristic of 'slow starters' in reading. It is not possible from this research to suggest just what the variables are that enable one group of initially poor readers to improve in reading performance over a 4-year period, compared with others who do not. The results do show that this improvement is *not* in any way associated with differences between the two groups on the Bannatyne WISC categories. This finding would support the conclusions of Henry and Wittman (1981) who compared the recategorised WISC scores of five groups of children with varying special learning needs and reported that Bannatyne's pattern was of little value in differentiating between students with learning disabilities and those without.

In the present study while the pattern of scoring on these categories for the failing readers was in accord with the pattern reported for dyslexics (Bannatyne, 1971) and disabled readers (Rugel, 1974a, 1974b; Smith *et al.*, 1977), it was of no more predictive utility than the score on the Sequential category itself.

The Bannatyne categories, therefore, while able to predict major differences in reading performance outcomes, were unable in this study to predict more subtle longitudinal changes in reading performance. This would suggest that the use of such categories is limited when making individual diagnosis and should only be considered in the context of the child's total learning environment and history.

REFERENCES

A.C.E.R (1972). *Primary Reading Survey Tests, Comprehension, Level B Form R.* Melbourne. Australian Council of Educational Research

Ackerman, P T., Peters, J E , and Dykman, R. A. (1971) Children with specific learning disabilities. WISC Profiles. *J. learn. Disabil.*, 4 (3), 150-166.

Bannatyne, A. (1968). Diagnosing learning disabilities and writing remedial prescriptions. *J. learn. Disabil.*, 1 (4), 28-35.

Bannatyne, A. (1971) *Language, Reading and Learning Disabilities.* Springfield, Ill.: Charles C. Thomas.

Bannatyne, A (1974). Diagnosis: a note on recategorization of the WISC scaled scores *J. learn. Disabil.*, 7 (5), 13-14

Congalton, A A. (1969). Appendix B: Status Ranking list of occupation in Australia. In Congalton, A. A , *Status and Prestige in Australia.* Melbourne. F W Cheshire Pub.

Dockrell, W. B. (1960) The use of Wechsler Intelligence Scale for Children in the diagnosis of retarded readers. *Alberta J. educ. Res* , 6, 86-91.

Dykman, R. A , Wallis, R. C., Suzuki, T., *et al.* (1970). Children with learning disabilities: conditioning, differentiation, and the effects of distraction. *Am. J. Orthopsychiat* , 40, 766-782

Dykman, R. A., Ackerman, P. T., Clements, S D and Peters, J. E. (1971). Specific learning disabilities: an attentional deficit syndrome In Myklebust, H. R (Ed.), *Progress in Learning Disabilities. (Vol 3)* New York. Grune and Stratton.

Egeland, B , Di Nello, M., and Carr, D (1970). The relationship of intelligence, visual-motor, psycholinguistic and reading readiness skills with achievement. *Educ. psychol Measur.*, 30, 451-458.

Fisher, R. A. (1949). *Statistical Methods for Research Workers.* (Tenth edition) Edinburgh Oliver and Boyd.

Fuller, G. B., and Friedrich, D. (1974/75). Three diagnostic patterns of reading disabilities. *Academic Therapy*, Vol. X (2). 219-231

Glasser, A. J , and Zimmerman, I L (1967). *Clinical Interpretation of the Wechsler Intelligence Scale for Children.* New York.: Grune and Stratton.

Henry, S. A., and Wittman, R. D. (1981). Diagnostic implication of Bannatyne's recategorized WISC-R scores for identifying learning disabled children. *J. learn Disabil.*, 14 (9), 517-520.

Huelsman, C. B. (1970). The WISC subtest syndrome for disabled readers. *Perceptual and Motor Skills*, 30, 535-550.

Hynd, G. W., Obrzut, J. E., Hynd, C. R., and Connor, R. (1978) Attentional deficits and word attributes preferred by learning disabled children in grades 2, 4 and 6. *Percept. mot Skills*, 47, 643-652

Hynd, G. W., Obrzut, J. E , Weed, W., and Hynd, C. R. (1979) Development of cerebral-dominance. dichotic listening asymmetry in normal and learning disabled children. *J. exp. Child Psychol.*, 28, 445-454.

Kaufman, A. S. (1975). Factor analysis of the WISC-R at 11 age levels between 6½ and 16½ years *J. consult. clin Psychol* , 43, 135-147.

Kaufman, A S. (1981). The WISC-R and learning disabilities assessment: state of the art. *J. learn. Disabil.*, 14 (9), 520-526

Levine, M., and Fuller, G. (1972). Psychological, neuropsychological and educational correlates of reading deficit. *J. learn. Disabil* , 5 (10), 563-571.

McKay, M. F. (1977). A four year longitudinal study of reading performance. Unpublished M.Ed. thesis, Monash University, Australia.

Neale, M. D. (1958). *Neale Analysis of Reading Ability.* London· Macmillan

Neale, M. D., McKay, M. F., and Thompson, G. B. (1979). The Neale Scales of Early Childhood Development and later reading performance: a multivariate analysis. *The Exceptional Child,* 26 (3), 133-143.

Neville, D. (1961). A comparison of the WISC patterns of male retarded and non-retarded readers. *J. educ. Res.*, 54, 195-197.

Obrzut, J. E., Hynd, G. W., Obrzut, A., and Pirozzolo, R. J. (1981). Effects of selective attention on cerebral asymmetries in normal and learning disabled children *Dev. Psychol* , 17, 118-125.

Roy, S. N., and Bargmann, R. E. (1958). Tests of multiple independence and the associated confidence bounds. *Ann math. Stat.*, 29, 491-503

Rugel, R. P. (1974a). WISC subtest scores of disabled readers· review with respect to Bannatyne's recategorization. *J. learn. Disabil.*, 7 (1), 48-55.

Rugel, R. P. (1974b). The factor structure of the WISC in two populations of disabled readers. *J. learn. Disabil.*, 7 (9), 581-585.

SATTLER, J M. (1974). *Assessment of Children's Intelligence.* (Revised reprint.) Philadelphia: W. B. Saunders.

SCHIFF, M. M., KAUFMAN, A. S., and KAUFMAN, N. L. (1981) Scatter analysis of WISC-R profiles for learning disabled children with superior intelligence. *J. learn Disabil.*, 14 (7), 400-404

SMITH, M. D., COLEMAN, J. M., DOKECKI, P. R., and DAVIS, E. E. (1977). Recategorized WISC-R scores for learning disabled children. *J. learn. Disabil*, 10 (7), 48-54.

WECHSLER, D (1949). *The Wechsler Intelligence Scale for Children.* New York: Psychological Corporation.

WECHSLER, D. (1974). *Manual for the Wechsler Intelligence Scale for Children — Revised.* New York: Psychological Corporation.

ZIMMERMAN, I. L., and WOO-SAM, J. (1972). Research with the Wechsler Intelligence Scale for Children. 1960-1970., *Psychology in the Schools*, 9, 232-271. (Special Monogr. Suppl.)

(*Manuscript received* 16*th April*, 1985)

Br. J. educ. Psychol., 55, 288-294, 1985

EVALUATING THE PICTURE FACILITATION EFFECT IN CHILDREN'S RECALL OF WRITTEN TEXTS

BY JENNIFER RUSTED* AND SANDRA HODGSON

(*Department of Psychology, Goldsmiths' College, University of London*)

SUMMARY. Forty nine-year-old children were required to read either factual or fictitious passages, presented with or without a relevant picture. Immediate free recall performance revealed a pictorial facilitation effect, the nature of which differed according to the type of passage. Pictures presented with story passages enhanced the recall of illustrated aspects of the text only, while those presented with factual passages enhanced the recall of both illustrated and unillustrated text content. This differential facilitation effect was related to the qualitatively different role that pictures play during encoding of the two types of texts.

INTRODUCTION

ILLUSTRATIONS have been used alongside children's reading material for many years, and numerous research studies have been completed in an attempt to identify the precise role of these pictures. In a review of this research, Levie and Lentz (1982) suggest that illustrations may improve performance in a number of ways, having attentional functions (i.e., directing attention to particular aspects of the material), affective functions (such as enhancing enjoyment and affecting attitudes), cognitive functions (facilitating comprehension and retention), and compensatory functions (whereby readers may use them to supplement their reading and understanding of the text). Under certain conditions, illustrations may also have a detrimental effect; for example in a study by Willows (1978) in which incongruent pictures were printed around a passage, intrusions from information contained in the pictures were recorded in the multiple choice comprehension questions. She suggested that pictures were distracting the readers; they spent less time reading the text and more time looking at the pictures, a consequence of which was a tendency to recall picture content rather than text content.

However, when pictures are congruent with the text, they have generally been found to enhance performance, particularly for textual information mirrored in the picture. 98 per cent of the experimental comparisons reviewed by Levie and Lentz reported that information represented both in the text and an accompanying picture was recalled better than when presented in an unillustrated text. For 85 per cent of these comparisons, the difference was statistically significant.

An issue which has not produced such consistent results concerns the nature and degree of facilitation observed for text content *not* represented in the accompanying picture. In children's recall of short stories, Vernon (1953) found that illustrations facilitated the learning of information which was directly illustrated in the pictures, but found no difference for unillustrated aspects of the text. She speculated that perhaps pictures laid undue emphasis on certain points and therefore distracted attention from the rest of the text. This hypothesis was subsequently supported in the work of Willows (1979, 1980) and of Peeck (1974) who reported no improvement in recall of story content not directly represented in the accompanying picture. Haring and Fry (1979) reported more specifically that when they presented children with fictitious narratives only recall for main idea units (i.e., those of high level importance in the story) benefited from picture representation. Even when included in the picture, lower level units were not recalled significantly better. Interestingly, studies using factual material have suggested that more global advantages may be observed with this type of text, although the number of studies is very limited. Of the four studies completed, two have observed improved recall for *all* aspects of text, whether or not they are explicitly illustrated in the accompanying picture (Rusted and Coltheart, V., 1979; Rusted and Coltheart, M., 1979). One study involving a much longer text (1000 Words) (Peeck, 1980) reported no advantage of illustrated presentation for unillustrated text content, and a fourth study (DeRose, 1976, cited in Levie and Lentz) found no improvement in recall for either illustrated or unillustrated text content.

---

* Now at Wayne State University, Detroit, Michigan

Rusted (1984) has presented a hypothesis which relates the source of the differences in picture facilitation of non-represented material to the type of material used (distinguishing between factual/expository texts and story texts) and the role of illustrations in directing encoding. She suggests that with expository texts pictures provide a conceptual framework for comprehension and encoding of the passage, producing a more elaborate and distinctive trace which is more readily accessible on demand than that produced by a similar unillustrated passage. The presence of the picture induces a qualitatively superior encoding and storage of the text *as a whole*, independent of the number of modalities in which individual aspects are represented. In contrast, encoding and recall of story texts are dictated by the children's use of the story schema (e.g., Mandler and DeForest, 1979). Rusted suggests that, because this schema provides an overlearned and highly effective strategy for organisation of the text at encoding, the alternative framework provided by the picture is overlooked. In consequence, the effect of illustrations is limited to the advantages obtained as a result of the dual modality presentation of specific aspects of the text. The picture serves to draw attention to the most important aspects of the text, to emphasise their significance and provide concrete visual images of that material. In effect, they encourage additional processing of these particular aspects which, under normal circumstances, will produce significant advantages at recall. Non-represented material will receive no additional processing, and recall availability will not be improved.

While there may be other explanations and contributing factors in the differential recall effects observed in earlier studies, the schema hypothesis provides an interesting and plausible explanation, and one which can be readily evaluated experimentally. The present study was completed as a direct test of this hypothesis. Information was presented to children either as a fictitious text or as a factual passage, and the effective use of identical illustrations was examined. According to the schema hypothesis, pictures provided with factual texts will enhance recall of both illustrated and unillustrated text content, while for fictitious presentation advantages will be observed for illustrated text content only.

METHOD

Subjects

Twenty boys and 20 girls from a middle school in Crawley, Sussex, were selected by their form teachers as being average or above average readers. Mean chronological age of the group was 9 years 7 months (range: 9 yr. 3m — 10 yr. 3m).

Materials

Two passages, one factual and one fictitious, were written about each of six unusual creatures; the mantella (a frog), the moloch (a lizard), the pitta (a bird), the saury (a fish), the tapir (an animal), and the weevil (an insect). Two practice passages were produced, one involving factual presentation and one involving story presentation, which followed the same format as the experimental passages but described a different creature (the kudu).

For each creature, the factual passage presented a short description, while the fictitious text provided a short story, relating an incident involving that creature. The two texts contained the same eight critical facts about the creature, four physical features which could be portrayed pictorially and four non-pictorial features concerning living/eating habits. All passages began with a statement naming the creature and its category; this was not part of the critical feature set. The critical features were presented in a similar order in both types of text, although with the fictitious passages some were inevitably more central to the story than others. Care was taken that pictorial and non-pictorial features provided the main idea units equally often across the six pairs of passages.

FIGURE 1

AN EXAMPLE OF THE FICTITIOUS AND FACTUAL PASSAGES USED

---

a. FACTUAL TEXT

The saury is a blue fish which swims near the surface of the water and often jumps above the waves. It is found in warm seas. It has long jaws and a row of small fins near its tail. It has a long thin body and swims very quickly.

b STORY TEXT

One day Saury the blue fish was swimming near the surface of the water and jumping above the waves at his home in the warm seas. Suddenly he became caught in a net. He set to work with his long jaws and freed the row of small fins near his tail. Then he cut a hole big enough to get his long thin body through and swam away quickly.

Factual passages were in general somewhat shorter than the fictitious ones (mean number of words being 45.8 and 66.5, respectively), as a result of the necessary 'filler' sentences required to produce a natural story style. For all passages, the only words which were unfamiliar to the children were the names of the creatures. A small pilot study ensured that within experimental conditions the six passages did not vary substantially in recallability. Post hoc examination of experimental data indicated a similar consistency across passages. An example of each type of text is shown in Figure 1.

Each passage was presented on a separate sheet of paper, written in standard handwriting using a 5mm letter stencil. The text was contained in the lower half of the sheet. A duplicate set of passages was produced, containing a coloured line drawing of the appropriate creature in the upper half of each page.

Design

Type of passage (factual/fictitious) was presented as a between subject variable, with 10 girls and 10 boys being randomly assigned to each condition. Each child received six experimental passages, three with pictures and three without in blocked presentation; half the subjects within each condition received illustrated passages first and half received unillustrated passages first. Presentation order of the six passages was randomised for each subject. In doing so, it was ensured that a given passage occurred equally often in the first and second blocks, and equally often with and without pictures for both the factual and the fictitious sets.

Recall protocols were scored in terms of the number of critical features recalled, for each passage a maximum of four pictorial features (features contained in both the text and in the picture, when present) and four non-pictorial features (features contained only in the text). Each child contributed four scores for analysis, which was completed on total recall scores across passages for illustrated pictorial, illustrated non-pictorial, unillustrated pictorial and unillustrated non-pictorial features. Earlier studies (Rusted and Coltheart, 1979) reported no sex differences in this task. In view of the small subject population in the present study, equal numbers of boys and girls were allocated to each experimental condition, but sex of subject was not included as a factor in the analysis of variance.

Procedure

The experiment was carried out during school hours. Subjects were tested individually in a quiet room. They were told that the study was about remembering, that they would be asked to read some short passages and afterwards to recall all they could about each. They were told that some of the passages would have pictures with them, and some would not. The session began with a practice passage, which was always presented without a picture, and which was the same type of text (factual or fictitious) as the child would receive in the experimental trials. The six experimental passages were then presented, in the predetermined order. Each child was asked to read the passages aloud, and reading errors (which were minimal) were corrected immediately. After each passage, the page was removed from view and the child was asked to recall all s/he could remember from the text. The child was cued with the name of the creature and its category. Recall was completed orally and monitored by the experimenter on a detailed scoring sheet. Neutral prompts (e.g., "Can you remember anything else?") were used to elicit full recall. No corrections were given for incorrectly recalled features. This procedure was repeated for each of the six passages. The duration of each individual session was approximately 15 minutes.

RESULTS

Recall performance was analysed using a three-way split plot ANOVA, with type of passage (factual or fictitious) as a between-subject factor, illustrations (present/absent) and type of feature (pictorial/non-pictorial) as within-subject factors. Mean recall scores are shown in Table 1, and graphically in Figure 2.

TABLE 1

MEAN RECALL SCORES FOLLOWING FACTUAL TEXTS OR STORY PASSAGES

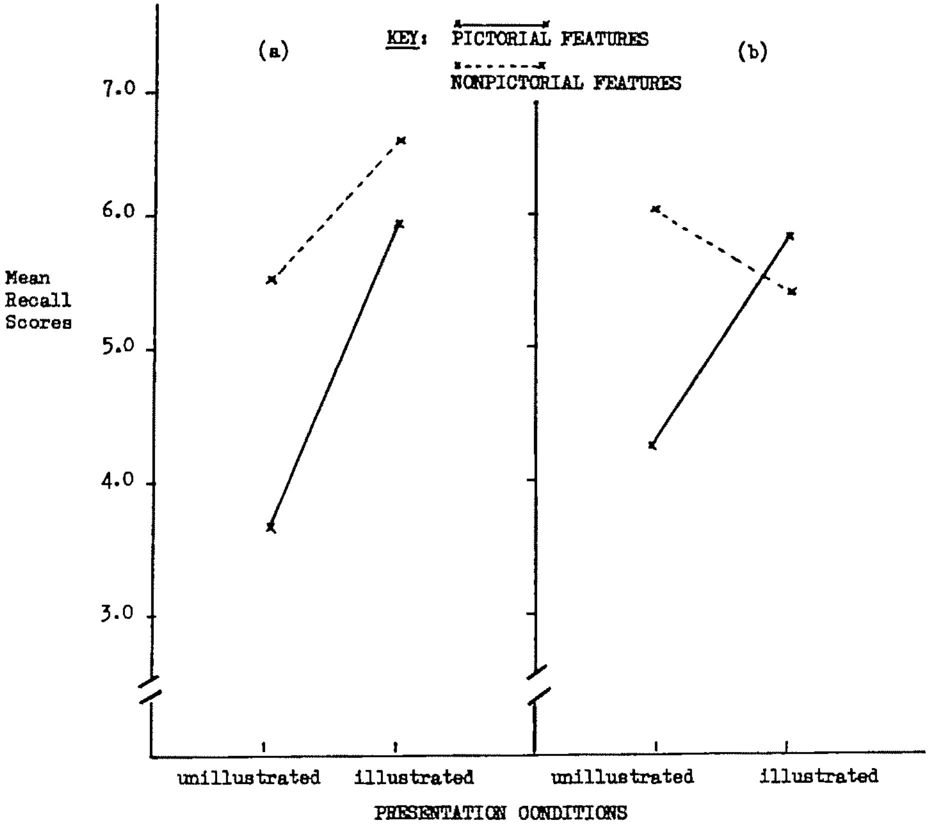
	Illustrated			Unillustrated		
	Pictorial Features	Non-pictorial Features	Total	Pictorial Features	Non-Pictorial Features	Total
Factual Text	5·95	6·60	12·55	3·65	5·50	9·15
Story Text	5·80	5·40	11·20	4·25	6·05	10·30

The analysis of variance revealed no significant difference in critical feature recall between the factual and the fictitious passages ($F < 1$). The presence of a relevant picture significantly improved recall ($F = 17.03$; df = 1, 38; $P < 0.001$), replicating the results of earlier studies, and there was also a main effect of type of features recalled ($F = 8.81$; df = 1, 38; $P < 0.01$), with more non-pictorial features being recalled than pictorial ones. There were two significant interactions in the analysis. Firstly, between type of passage and presence/absence of illustrations ($F = 5.76$; df = 1, 38; $P < 0.03$); as can be seen in Table 1, the illustrations had greater effect on recall when presented with the factual texts. Tests of simple main effects indicated that the advantage of illustrated over unillustrated texts attained the 0.05 significance level for factual passages only. The other significant interaction occurred between presence/absence of illustrations and pictorial/non-pictorial feature recall ($F = 6.8$; df = 1, 38; $P < 0.02$); tests of simple main effects indicated that the presence of illustrations significantly improved recall of pictorial facts, but not of non-pictorial facts.

Although the three-way interaction between passage type, illustrations and

FIGURE 2

MEAN RECALL PERFORMANCE FOLLOWING (a) FACTUAL AND (b) STORY TEXTS, ACCORDING TO PRESENTATION CONDITIONS AND FEATURE TYPE



feature type was not significant, planned comparisons using t-tests were completed to evaluate the predicted differential facilitation effects across the two types of passages. The results were as follows. For factual passages, illustrations significantly improved recall of both pictorial ($t=4.94$, $df=19$, : $P<0.01$) and non-pictorial ($t=1.96$, $df=19$, $P<0.05$) facts. For fictitious passages, illustrations significantly improved recall for pictorial facts ($t=2.21$, $df=19$, $P<0.05$) but not for non-pictorial facts ($t=0.93$), which in fact tended to be recalled better in the absence of illustrations. This reversal may have contributed to the non-significant three-way interaction, and to the pattern of effects observed in the interactions reported above.

DISCUSSION

The picture facilitation effect observed in the present study was qualified by a number of interactions. Despite the significant main effect of illustrations, the advantage of illustrated over unillustrated presentation was greater for factual/expository texts than for stories, and this was confirmed in the analysis of the two-way interaction between presence/absence of pictures and type of text, in which the facilitation effect was shown to be statistically significant only for factual texts. In addition, the two-way interaction between presence/absence of pictures and type of facts recalled suggested that the picture facilitation effect was limited to pictorial features of the text. Both of these interactions appear to be linked to the decrease in recall of non-pictorial facts occurring for illustrated over unillustrated story texts, which masks the observed pattern of facilitation effects more accurately described by the results of the *a priori* comparisons.

As indicated in these comparisons, picture facilitation effects were observed for both pictorial and non-pictorial features contained in the factual texts, but were limited to pictorial features in the story texts. In story recall, an advantage of *un*illustrated over illustrated presentation conditions was apparent for non-pictorial text content, although this failed to reach the 0.05 level of significance. This latter finding is entirely consistent with the conclusions of earlier researchers (e.g., Vernon, 1953; Peeck, 1974; Willows, 1979, 1980) that pictures may draw attention away from the non-illustrated text content, and encourage undue emphasis on those aspects of the text portrayed in the picture. A major determinant of story recall in children and adults is the structured presentation format known as the story schema (Mandler and DeForest, 1979; Nezworski *et al.*, 1982), whereby the story is unfolded in a well-established order which can be used as an organisational aid both at encoding and retrieval. From experience, the general expectation of the reader is that the picture (a) is relevant to the story line and (b) emphasises those aspects of the text considered most important by the writer (or in this case, the experimenter). Considerable importance is therefore attached to the accurate encoding of picture content. In this instance, a picture which described the main character but provided no event description improved recall of the visually represented features, but at the expense of event recall which included both pictorial and non-pictorial text content. In the absence of a picture, this bias was removed, encoding and recall reflected the hierarchical organisation of content according to the story structure, and non-pictorial feature recall was no longer disadvantaged. The limited role of pictures in story recall is clearly demonstrated in the pattern of results reported here, which suggest the need for careful and considered selection of pictures both in experimental studies and, at a more general level, in the production of story books for children. In line with previous studies, the picture content appears far less critical for factual texts. It is not assumed to provide clues as to the importance of specific aspects of the text, and the advantage of illustrated presentation is not limited to improved recall of the pictured features. Rather, the reader seems to make effective use of the picture as a concrete base for efficient encoding of *all* the associated material, promoting a more distinctive and accessible encoding of the text as a whole.

There are obvious difficulties in attempting a direct comparison between recall scores from the fictitious and factual texts, and this study has a number of attendant problems.

Concentrating on a critical feature set meant that in general the recall performance for story passages included information not subsequently credited to the overall score. While care was taken in text construction to limit the amount of non-critical "filler" features, the need to produce plausible story lines necessarily dictated the degree of success. The story line also determined the presentation order of the feature set. As a control, this order was mirrored in the factual text, although it sometimes led to a less than natural sequence in this text.

Accurate recall of text content was stressed to all subjects during the preliminary instructions and following presentation of the example passages. Recall protocols indicated that the children complied with these instructions; critical feature recall was not artificially reduced by introduction of uncredited 'plausible substitutions' or interpretations. This was a more serious potential problem in the story texts. However, the length of the texts, being well within the scope of most of the readers, did not induce such alterations under immediate recall conditions. Synonymous substitutions in either factual or fictitious recall were credited.

The present study did not control for possible interactions with idea importance within the story texts, although this is certainly an issue which warrants more detailed examination. The texts used here were deliberately written as short, rather low-key stories, in which for the most part no particular feature was given undue prominence. Although it may be argued that this in fact limits the usefulness of the story schema as an aid to recall, the story format would nevertheless be expected to promote sequential recall of 'events' in the same way as the story schema promotes sequential recall of episodes. The length, cohesiveness and lack of redundancy, in the texts used here would favour a high level of performance with such a strategy. Story recall might therefore be expected to be superior to the relatively unstructured factual recall, and also to be less susceptible to the effect of picture inclusion. This was not the case. While the subject matter and, more importantly, the size of the texts appear in this instance to be entirely compatible with our manipulations, we acknowledge the need to explore the robustness of the findings across a much wider range of materials and in particular the discourse structure more typical of the narratives from which the story schemata were originally formulated.

In conclusion, the results reported here reinforce previous findings and confirm the consistency of the picture facilitation effect in children's recall of written texts. More importantly, they provide support for the hypothesis that the nature of this improvement differs qualitatively according to the type of material presented. The role of pictures presented with story texts appears to be limited by the readers' reliance on the story schema to direct encoding, and the expectation of salient feature representation in the accompanying picture. In consequence, enhanced recall occurs only for those aspects of the text accurately represented in the picture. The advantages observed for illustrated presentation of factual texts reflect global improvements in text recall, suggesting a role in the promotion of distinctive encoding rather than the content-specific advantages found with story texts.

Reprints available from J. Rusted, Deptartment of Psychology, Wayne State University, 71, W. Warren Ave., Detroit, Michigan 48202.

REFERENCES

Haring, M. J , and Fry, M A. (1979). Effects of pictures on children's comprehension of written text *Educ Commun. Technol* , 27, 185-190.

Levie, W. H., and Lentz, R. (1982). Effects of text illustrations: A review of research. *Educ. Commun Technol* , 30, 195-232.

Mandler, J M., and DeForest, M (1979) Is there more than one way to recall a story? *Child Dev.*, 50, 886-889

Nezworski, T., Stein, N. L , and Trabasso, T. (1982). Story structure versus content in children's recall. *J. verb. Learn verb. Behav* , 21, 196-206.

Peeck, J (1974) Retention of pictorial and verbal content of a text with illustrations. *J. educ Psychol* , 66, 880-888.

Peeck, J (1980) Experimenter-provided and learner-generated pictures in learning from text Paper presented at the Annual Meeting of the American Educational Research Association, Boston.

Rusted, J. M. (1984). Differential facilitation by pictures of children's retention of written texts. a review *Current Psychological Research and Reviews*, (in press).

Rusted, J M., and Coltheart, M. (1979) Facilitation of children's prose recall by the presence of pictures. *Memory and Cognition*, 7, 354-359

Rusted, J M , and Coltheart, V (1979) The effect of pictures on the retention of novel words and prose passages. *J. exp. Child Psychol* , 28, 516-524.

Vernon, M. D. (1953) The value of pictorial illustrations. *Br. J. educ. Psychol.*, 23, 180-187.

Willows, D M. (1978) Individual differences in distraction by pictures in a reading situation. *J. educ. Psychol.*, 70, 837-847.

Willows, D. M. (1979). Reading comprehension of illustrated and nonillustrated aspects of text. Extended version of paper presented at the Annual Meeting of the American Educational Research Association, San Francisco

Willows, D M. (1980). Effects of picture salience on reading comprehension of illustrated and nonillustrated aspects of text. Paper presented at the Annual Meeting of the American Educational Research Association, Boston

(*Manuscript received* 20*th February*, 1985)

Br. J. educ. Psychol., 55, 295-303, 1985

THE EFFECTS OF TEACHER PRESENTATION RATE AND PRAISE ON LD STUDENTS' ORAL READING PERFORMANCE

BY CRAIG DARCH
(*Auburn University, Alabama*)

AND RUSSELL GERSTEN
(*University of Oregon*)

SUMMARY. Four learning disabled students were taught beginning reading with a Direct Instruction programme. The teacher systematically varied the rate of instructional presentation (rapid pace *vs.* slow pace) and frequency of praising (praise *vs.* no praise). A modified reversal design was used to evaluate both the isolated and interactive effects of the individual instructional components. Experimentation lasted 25 school days. The results indicated that increases in teacher presentation rate and increased use of praise lead to improvement in levels of on-task behaviour and correct student responding. Each of the variables in isolation has an effect. Praise appears to be more powerful than presentation rate. The combination of the two variables has the most powerful effect. These findings extend the research of earlier studies to learning disabled students.

INTRODUCTION

NATURALISTIC studies of classrooms with learning disabled students (Leinhardt *et al.*, 1981) and other low performing students in remedial programmes (Stallings, 1980; Gersten *et al.*, 1982) have shown that certain sets of teacher behaviours are consistently correlated with growth in reading — particularly oral reading for the very low performing students. One of the most exciting implications of these findings is that most of the performance variables specified in the aforementioned studies — praise rate, provision of academically oriented feedback to students, allocation of time for oral reading — appear reasonably easy to alter. In fact, Stallings (1980) demonstrated that, when high school remedial reading teachers were observed, given objective feedback on their performance, and in-service training and practice on the desired behavioural change, significant improvements were found in both teacher performance, and most importantly, student gains in reading.

The present study attempts to do much the same thing, using only one group of students. With the modified reversal design (Hersen and Barlow, 1976) used in this study, one can isolate the specific effects of each aspect of teacher behaviour on student performance.

This study attempted to merge research with the teacher training experience as Tawney (in press), among others, has advocated. The experimental teacher was a student teacher, completing her graduate work in special education; the master teacher (assisted by the supervisor) designed the study, collected the data, and shared the results with the student teacher. Thus, the trainee was able to see the impact of changes in her behaviour on her students' performance.

The two teacher performance variables selected for this study were (a) teacher presentation rate, and (b) use of praise. Both have been found to be related to increased growth in reading — presentation rate in the Gersten *et al.* (1982) study, praise in the Leinhardt *et al.* (1981), and Stalling's (1980) studies. The research on praise and verbal reinforcement is myriad (see for example, Sulzer-Azaroff and Mayer, 1977; Walker 1979, for reviews of research and its use with mildly handicapped students). Verbal praise tends to increase young students' attention to the task in hand; often, but not always, student accuracy rate also increases (cf. Carnine, 1978; Marcy, 1979, for cases where on-task behaviour increases, but academic performance remains the same).

On the other hand, virtually no research has been done on the effects of teacher presentation rate on student learning. Using a reversal design, Carnine (1976) demonstrated that use of a rapid presentation rate (often called pacing) was accompanied by an increase in on-task behaviour by both subjects in the study. One of the two subjects demonstrated increases in the per cent of correct responses and participation; effects on these measures were erratic for the second subject.

A second study by Carnine (1981) demonstrated that when teachers simultaneously (a) increased presentation rate, (b) increased praise rate, (c) provided clear signals, and (d) gave consistent, immediate corrections of student errors, the proportions of correct student responses of on-task behaviour increased in the three preschoolers. Neither of these studies involved handicapped students; in both cases, the subjects were low performing students in regular classroom settings. Both, however, suggest that praise and pacing might improve the on-task behaviour and accuracy rate of mildly handicapped students, especially since the boundaries between learning disabled and low performing students tend to be quite amorphous (Shepard *et al.*, 1983).

Only two studies using rapid pacing and praise as independent variables have been conducted with special education populations. In a study involving three autistic children, Koegel *et al* (1980) demonstrated that decreases in inter-trial intervals (i.e., increase in teacher presentation rate) led to increases in the proportion of correct student responses and seemed to enhance acquisition. West (1981) examined the impact of pacing and praise on the performance of five students referred to a clinic for extreme rates of disruptive behaviour. He found that rapid pacing led to significant decreases in proportion of disruptive behaviour. The mean for the rapidly paced sessions was 55 per cent, for the slowly paced sessions, 78.5 per cent. There was no discernible effect for praise rate. The results on performance accuracy are very interesting. Unlike in the Carnine studies, student accuracy rate was actually lower with the rapidly paced lessons (79 per cent versus 86 per cent). Again, verbal praise had no impact on this measure. The reader should note the extremely high rates of disruptive behaviour of the students involved in the study. The impact of verbal praise and presentation rate may be quite different for students with serious learning problems but with much lower rates of inappropriate behaviour.

The present study was designed to extend the research on rapid pacing and feedback to learning disabled students. Also, both the isolated and inter-active effects of rapid pacing and praise were evaluated.

METHOD

Children and setting

The children for this study were four students, three boys who ranged in age from 7.8 to 8.6 years and one girl, age 8.3 years. Each was identified as learning disabled by school psychologists whose assessment included history of difficulty in basic word decoding. The mean IQ for the four children was 101 (SD = 4·39). Each child met state guidelines for placement in a learning disabilities classroom; this included normal potential, as measured by the IQ scores, and a significant discrepancy between expected and actual achievement levels. Each child in this study had failed to learn basic reading skills associated with a first grade curriculum and was reading well below a first grade achievement level. The school psychologist indicated that these children exhibited a severe learning disability. At the beginning of this study, each child had just been placed into the learning disabilities classroom. Students' individualised educational programmes (IEPs) included learning basic sound/symbol relationships and blending sounds into words which was indicative of their low reading performance.

The reading class met daily for a 45-minute period. The first 20 minutes were used for the experimental intervention. The remainder of the class period was devoted to the students working on individually designed worksheets. The four students were taught as a group. The teacher for this study was a 23-year-old woman who was completing a graduate practicum in learning disabilities.

Reading programme

The Distar Reading I programme (Engelmann and Bruner, 1975) was used for this study. In this programme, the rate of teacher presentation (pacing) is expected to be rapid (approximately ten tasks per minute).

Before the study began, the teacher and experimenter practised providing signals for student responding, contingent praise, and increasing the rate of instruction. Training consisted of demonstration, practice, and feedback. These sessions occurred after school and lasted approximately 30 minutes for five consecutive days.

Experimental design

A modified reversal design, with replication across subjects, was used (Hersen and Barlow, 1976). The advantage of this particular design is that it allows demonstration of both the isolated and interactive effects of individual intervention components. Because most instructional systems are multi-component, the ability to identify those aspects of a treatment that contribute most to increased student performance is very important. Data were collected daily during this study and experimentation lasted a total of 25 school days.

Baseline 1

During this phase, the teacher presented material from the Distar lessons by closely following the daily lesson scripts, with two exceptions. Although the teacher told students whether their responses were correct or incorrect, she did not verbally praise correct answers or instances of appropriate social behaviour. If the group gave a correct response, the teacher would merely tell the students the answer was correct, and move on to the the next step in the lesson. If the students gave an incorrect answer, the teacher would model the correct response and proceed to the next part of the lesson.

The rate of teacher presentation of the instructional content was also controlled. To ensure a slowly paced instructional delivery, the teacher was instructed to pause 5 seconds (by counting to herself) between the end of a student's response and the presentation of the next part of instruction. This procedure was similar to the one used by Carnine (1976). Baseline 1 lasted for four days.

Rapid Pace 1

This phase lasted four days and was exactly the same as Baseline 1 except that the pace of instructional presentation was increased. Instead of creating a 5-second pause between the completion of one task and the introduction of the next, the teacher was instructed to proceed immediately to the next part of the lesson. It must be emphasised that increased pacing had nothing to do with the rate at which the teacher spoke. Instead, pacing was defined as the time between task presentations.

During this phase, the teacher gave only informational feedback to the student; she did not verbally praise their behaviour. Corrections were done in the same manner as in Baseline 1.

Rapid Pace and Praise 1

This phase replicated the rapid pacing procedure and added praise for appropriate social behaviour and correct academic responses. During each instructional presentation, the teacher identified instances of appropriate behaviour and praised students, using short specific statements. The teacher reinforced behaviour such as eyes focused on the text, keeping hands and feet to oneself, staying in one's seat, and responding to the teacher's signal. Also, the teacher would praise one or more of the group for responding correctly. This phase lasted five days.

Praise Only

Next, rapid pacing was dropped to demonstrate the differential effectiveness of praise *vs.* pacing. The teacher again presented material at a slow pace, using the 5-second pause technique. The teacher replicated the verbal reinforcement procedure described in the previous condition. This phase lasted three days.

Baseline 2

This phase was identical to Baseline 1 and lasted three days.

Rapid Pace 2

This phase replicated Rapid Pace 1 and lasted three days.

Rapid Pace and Praise 2

This phase, a replication of Rapid Pace and Praise 1, was important in again showing the effects of the two variables in combination. It lasted three days.

Dependent variables

Per cent Correct. Per cent Correct was calculated by dividing the number of opportunities available to respond to a teacher question or direction by the number of correct answers. If the student did not correctly respond within one second of the teacher's signal, the response was marked as incorrect. In order for a student's response to be considered correct, the student had to voice the answer loudly enough to be heard by the observer, who was seated approximately four feet away. Data were taken during both individual turns and group responses. Procedures are discussed below.

Per cent On-Task Behaviour. On-Task was defined as eyes on the teacher or the presentation book, keeping hands and feet to oneself, and responding on time to the signal. Behaviours such as talking out, talking to classmates, playing with objects, out-of-seat without permission, or fighting were classified as off-task.

Data collection procedure

Three university graduate students served as primary data collectors. One collected data on on-task behaviour while the other recorded correct responses. On alternating days, the third graduate student would observe and record either on-task behaviour or correct responses to assess inter-observer reliability. Data were collected on the four subjects in the following manner. Observers who coded on-task bahaviour used a 10-second time-sampling technique; they would observe the student for 9 seconds and record the student's behaviour during the tenth second. If the student was engaged in appropriate behaviour for the entire nine seconds, the observers would record a slash (/) in the appropriate interval on the data sheet. If the student was engaged in any inappropriate behaviour during the nine seconds, the observers recorded a 0 in the specific interval. A different student was coded each 10-second interval. The observers systematically coded the four subjects in a predetermined order; therefore both observers were recording the same subject at the end of each 10-

second interval. One complete rotation through the subjects took 40 seconds and the process was repeated continuously for the duration of the session.

The technique to record correct responding followed the procedures described by Carnine (1976). Each subject was recorded individually for five consecutive tasks, even though all subjects responded to each task. A task was defined as a teacher presentation of either a sound or a word. For example, each task required the students to: (a) give the sound that a letter or letter combination made, (b) sound out a word, (c) to read the word as a whole unit. This cycle was repeated numerous times during the 20-minute lesson.

RESULTS AND DISCUSSION

Inter-observer agreements for on-task behaviour ranged from 80 to 95 per cent with a mean of 90 per cent; for percentages of correct answering, they ranged from 88 to 100 per cent with a mean of 95 per cent.

Accuracy

The results on the percentage of correct responses for each subject are shown in Figure 1. Each student was performing well below acceptable level during the Baseline 1 condition. Each student's performance was affected by the introduction of each phase of the experimental intervention. With the introduction of the first component of the intervention, Rapid Pace 1, each student demonstrated an increase in accuracy. This increase in performance was replicated in the Rapid Pace 2 condition. This consistent effect shows that teachers of LD students can improve academic performance merely by increasing the rate of instructional presentation. This finding replicates Carnine's (1976) study. This result also supports the findings of a correlational study conducted by Gersten *et al* (1982), in which the students of teachers who consistently paced their lessons briskly gained more in reading achievement than their peers. It is important to note that this increase in pacing was not due to the teachers speaking more quickly. Rather, the pauses between each segment of the lesson were reduced.

The combination of Rapid Pacing and Praise led to an even stronger effect. Students 1 and 2 improved from a baseline accuracy of 40 to 86 and 79 per cent respectively. Student 3's growth was not as dramatic (from 53 to 65 per cent). Student 4 was the lowest performer in the group. Though she demonstrated some growth in accuracy (18 to 31 per cent), she continued to perform well below acceptable levels.

When the Praise Only condition was administered, performance gains on correct responses are maintained in student 3, while slight decreases are shown in the three other students.

Levels of performance during the last three phases replicate the results in the first three conditions of the experiment. The finding that the combination of Rapid Pacing and Praise can produce powerful increases in performance of LD students is an important finding for the classroom teacher.

On-Task

The results for on-task behaviour for each subject are shown in Figure 2. The findings replicate those for correct responding. Rapid Pacing and Praise both contributed to increased on-task behaviour. These findings have implications for teachers whose students have high levels of off-task behaviour.

The pattern of changes in the percentage of on-task behaviour across conditions was similar to the pattern observed for correct responses. Subject 1 spent 52 per cent time on-task during B_1. The introduction of the Rapid Pace condition produced a clear increase in mean on-task behaviour to 60 per cent (range 59 to 63). This further

FIGURE 1

PERCENT CORRECT RESPONDING OF EACH SUBJECT DURING ALL PHASES OF EXPERIMENTATION

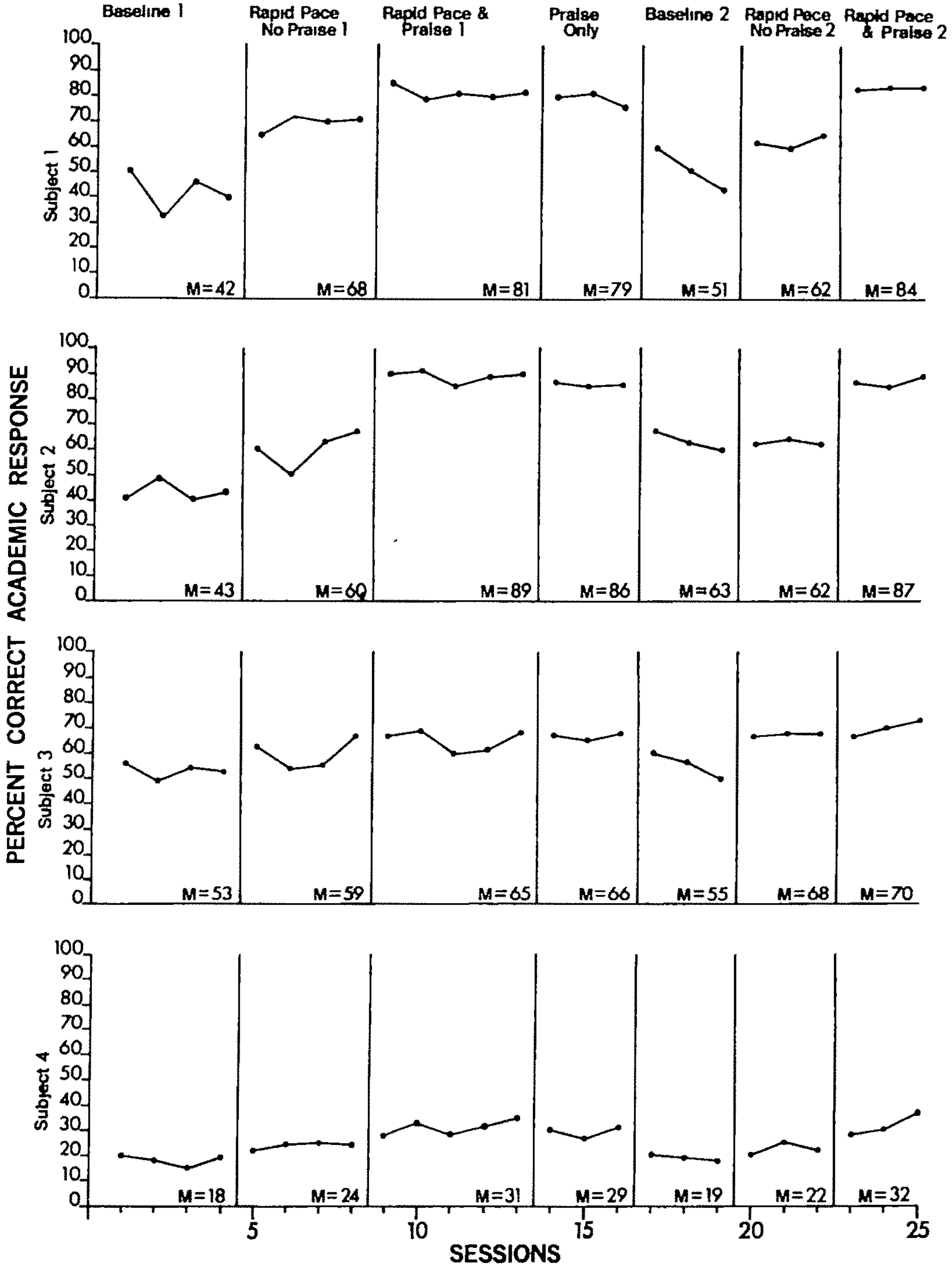
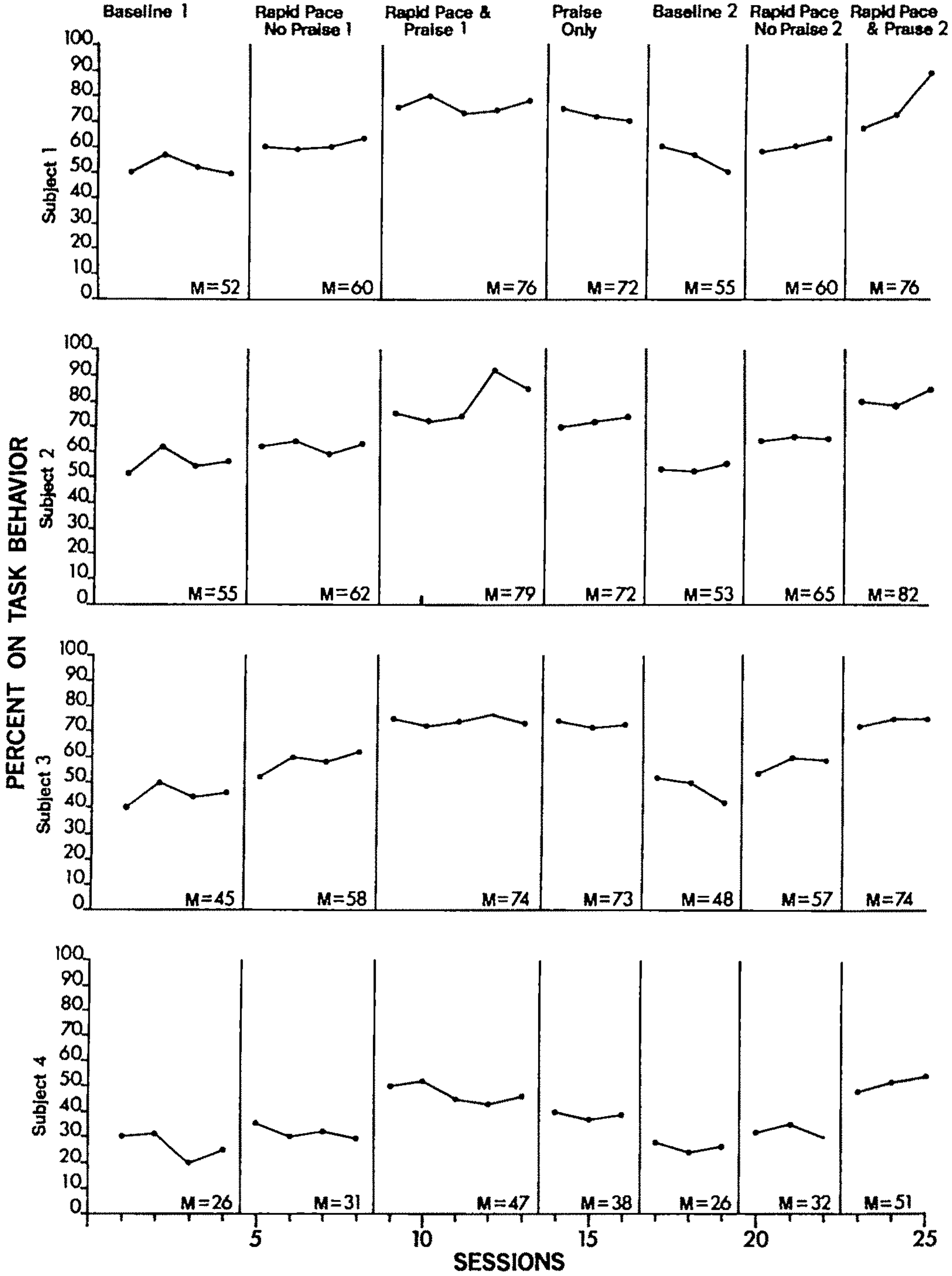


FIGURE 2

PERCENT ON-TASK BEHAVIOUR OF EACH SUBJECT DURING ALL PHASES OF EXPERIMENTATION



improved (to a mean of 76 per cent) during the combination Rapid Pace and Praise condition. The Praise Only phase is only slightly lower (mean of 72 per cent). The reversal to B_2 leads to a dramatic decrease, a mean of 55 per cent. The reintroduction of Rapid Pacing again leads to a slight increase (mean of 60). The reintroduction of Rapid Pace and Praise 2 replicates the previous effect (mean of 76 per cent).

Subject 2 demonstrated a similar pattern of on-task behaviour across experimental conditions: a slight increase with Rapid Pace only, a somewhat dramatic increase with the combination of the two variables (79 per cent), and a slight drop when praise only is used (72 per cent). When both variables are removed (Baseline 2), performance drops to the original level (range of 52 to 54 per cent). The introduction of the Rapid Pace 2 phase saw on-task increase to 65 per cent, while the Rapid Pace and Praise 2 further increased on-task behaviour to an average of 82 per cent.

The performance of Subject 3 is similar to the others. In his case, however, there is virtually no drop from Rapid Pace and Praise to Praise only.

Subject 4 had a mean on-task level of 26 per cent, clearly the lowest level of all the subjects. When Rapid Pace 1 was introduced, mean percent on-task behaviour rose to 31 per cent. When the Rapid Pace and Praise condition was introduced this subject's on-task behaviour increased to 47 per cent. This is almost double the baseline level. Her on-task behaviour drops to 38 per cent during the Praise only condition; this is more of a drop than the other subjects. The return to baseline (B_2) produces a dramatic drop (back to a mean of 26 per cent). Positive effects are replicated for Rapid Pace 2 and the Combination Condition (51 per cent on-task). Clearly, more powerful intervention procedures are warranted for this subject.

CONCLUSIONS

These results indicate that increases in teacher presentation rate and increased use of praise lead to increases in on-task behaviour and to relatively large increases in student accuracy for two of the four LD students in the group, and smaller, but appreciable gains for the other two students. Each of the variables in isolation has an effect. Praise appears to have more of an impact on this group of students than presentation rate. The combination of the two variables has the most powerful impact.

These results are consistent with Carnine's (1976) findings with non-handicapped beginning readers. They are quite different from West's (1981) findings with similar age children exhibiting high levels of disruptive behaviour. West found that rapid pacing was more effective than praise in decreasing off-task behaviour, but that the accuracy rate actually *decreased* during the rapid pacing phases. The students in the West study were quite different from those in the current study, exhibiting much higher rates of disruptive behaviour. It is promising to see that the intervention had an impact on both student accuracy and on-task behaviour rate. Several studies have shown that verbal praise can increase on-task behaviour without any concomitant increase in the amount the student learns (Marcy, 1979; Carnine, 1978; West, 1981). A large-scale study by Fisher *et al.* (1980) demonstrated that both high accuracy rates (e.g., 80 to 90 per cent) and high engagement rates are necessary for maximum academic gains.

Teachers must look to both instructional modifications and use of consequences when developing instructional programmes for skill deficient students. The powerful effects of pacing and praising can be important in allowing students to succeed in the early stages of a remedial education programme. The finding that rapid pacing alone increases performance might be helpful as a method to eliminate (or reduce), at some point in a student's instructional programme, supplemental reinforcement programmes, (i.e., token systems, high rates of praise). A number of investigators have suggested that use of extrinsic reward systems may produce students who only learn to earn tokens. These students may fail to develop intrinsic motivation for academic tasks. If teachers modify instructional delivery variables (i.e., pacing) some students may not need supplemental reinforcement programmes. Further, if a student has already been placed on a point system, an increased instructional pace may facilitate early removal of this intervention.

Future research should look at other teacher performance variables that correlation research suggests are effective. Both Leinhardt *et al.* (1981), Stallings (1980), and the study by Gersten *et al.* (1982) found that when teachers provide supportive, but academically oriented feedback to students after incorrect responses, learning is increased. Yet, to our knowledge, no research has been done to examine experimentally the impact of this variable on student performance. This study did demonstrate that application of two relatively "easy to master" teaching skills can increase how much students learn. Single subject designs can allow for an integration of research into the teacher training situation, and for ongoing demonstrations to student teachers of the impact of their teaching on the students they work with.

REFERENCES

CARNINE, D. W (1976). Effects of two teacher-presentation rates on off-task behavior, answering correctly, and participation. *J appl. Behav. Analysis*, 9, 199-206.

CARNINE, D W. (1978). High and low implementation of direct instruction teaching techniques *Educ and Treatment of Children*, 4, 43-51.

CARNINE, D. W. (1981). High and low implementation of direct instruction teaching techniques. *Education and Treatment of Children*, 4, 43-51.

CARNINE, D W., and FINK, W T. (1978). The effects of eye contact on attending and correct responding during small group instruction. In BECKER, W. C., and ENGELMANN, S , *Technical Report* 78-1 Eugene, Oregon· University of Oregon Follow Through Project.

ENGELMANN, S E., and BRUNER, E. (1975). *Distar Reading I.* (Second edit.). Chicago: Science Research Associates.

FISHER, C. W , BERLINER, D. C., PHILBY, N W., MURLIAVE, R., CAHEN, L. S., and DISHAW, M M. (1980). In DENHAM, C., and LIEBERMAN, A. (Eds), *Time to Learn*, pp 7-22 Washington: NIE.

GERSTEN, R., CARNINE, D., and WILLIAMS, P. (1982). Measuring implementation of a structured educational approach an observational approach *Educ. Eval. Policy Analysis*, 4, 67-79.

HERSEN, M , and BARLOW, D. H. (1976). *Single Case Experimental Design.* New York: Pergamon Press.

KOEGEL, R L., DUNLOP, G., and DYER, K. (1980). Intertrial interval duration and learning in autistic children *J appl. Behav Analysis*, 13, 91-99.

LEINHART, G., ZIGMOND, N , and COOLEY, W. W. (1981). Reading instruction and its effects. *Am. educ. Res. J.*, 18, 343-361.

MARCY, C R. (1979) Effects of variable interval schedule of praise on attending and computation proficiency of third graders Unpublished Doctoral Dissertation, Eugene, Oregon.

SHEPARD, L. A , SMITH, M. L., and VOJIR, C. P. (1983). Characteristics of pupils identified as learning disabled. *Am educ Res J* , 20, 309-331.

STALLINGS, J A (1980). Allocated academic learning time, or beyond time on-task. *Educ. Res.*, 9, 11-16

SULZER-AZAROFF, B , and MAYER, G. R. (1977) *Applying Behavior Analysis Procedures with Children and Youth.* New York: Holt, Rinehart and Winston.

TAWNEY, J W Empirical verification of instruction a realistic goal? In HEWARD, H , HERON, T , HILL, D. S , and TRAP-PORTER, J. (Eds), *Focus on Behavior Analysis in Education* Columbus, Ohio Charles Merrill (in press)

WALKER, H. M. (1979). *The Acting-out Child· Coping with Classroom Disruption* Boston: Allyn and Bacon.

WEST, R. P. (1981) The effect of two levels of teacher presentation rate and reinforcement rate on classroom disruptions, performance accuracy, and response rate Unpublished Doctoral Dissertation. Salt Lake City, Utah.

(*Manuscript received* 3rd *May*, 1985)

RESEARCH NOTES

Br. J. educ. Psychol., 55, 304-309, 1985

PAIRED READING VERSUS RELAXED READING: A COMPARISON

BY GEOFF LINDSAY
(*Sheffield Psychological Service, and University of Sheffield Division of Education*)

ALISON EVANS
(*Derbyshire Psychological Service*)

AND BEN JONES
(*University of Sheffield Division of Education*)

SUMMARY. This study compared the effectiveness of two methods of parental involvement in the development of their children's reading Each method was adopted with a group of 10 children who were sub-divided with respect to whether their progress during the 6-week project was monitored by home visit or telephone. The results suggest that both methods can lead to substantial improvement in reading ability. The failure to find a significant difference between methods suggests that the simpler and more cost-effective approach, Relaxed Reading, is to be preferred

INTRODUCTION

The involvement of parents in the development of their children's reading has been, and still is, a contentious area. At an extreme is the view that reading is a complex skill and therefore should only be taught by skilled professionals, but evidence from biographies and anecdotes suggests that many children have learned to read from their parents or other non-professionals, and even that some have taught themselves. During the recent decade, therefore, there has been an increasing interest in the harnessing of parental involvement as a resource.

The studies by Hewison and Tizard (1980) and Tizard *et al* (1982) provided evidence that parents can indeed be very important in the development of their children's reading, and that schools can work collaboratively with them. The latter study made use of a relatively simple technique whereby children were encouraged to read at home to their parents, using books provided by the school. These children made significantly greater gains in reading scores than control children who either received extra help in school from a teacher, or who received no extra involvement. A longer term study at the Belfield Community School in Rochdale (Jackson and Hannon, 1981; Hannon *et al.*, 1985) has shown that this type of project can maintain the involvement of parents and children over 2 to 3 years.

Much interest has also developed in a more structured technique of parental involvement known as Paired Reading. This was described by Morgan and Lyon (1979), and later by Bushell *et al* (1982). Since that time there has been a large number of small scale studies describing its effectiveness (e.g , Topping and McKnight, 1984). In this approach, children and parents are taught specific techniques of Simultaneous and Independent Reading. With the former the child and parent read together in synchrony. This, it is argued, ensures that the child does not meet with failures as, if unsure of the word, the child can read it immediately after the parent. In addition the parent is taught to be reinforcing by being fulsome with praise. When the child moves on to Independent Reading, he or she will read alone until encountering a word about which they are unsure. A simple signal leads the parent to join in, i.e., Simultaneous Reading restarts, and this continues until the child indicates a desire to read alone once more. (Fuller details are provided in the references cited above.)

There is a good deal of evidence which suggests that children who engage in Paired Reading schemes, which commonly last for a matter of a few months, make significant gains in reading scores. Topping and McKnight (1984), for example, in a review of 11 studies report gains of between 1.75 and 9 times those to be expected on the basis of the children's previous development. However, these studies used each child as its own control and did not compare other types of intervention.

The question arises whether the benefits are due specifically to the types of teaching approaches, or to other factors including a general reduction in anxiety which has been noted in these and other studies. The implications of such a distinction have practical as well as theoretical relevance as Paired Reading projects normally include costly home visiting schemes to check that the parent and children are performing Simultaneous and Independent Reading correctly.

The present study comprises a comparison of the effects of a Paired Reading intervention with those of a simpler method which we call Relaxed Reading. This method has as its central aim the development of children reading with their parents in a way which is non-anxiety provoking. Parents and children are shown how such a process can be done badly, with the parent being critical and negative and the child becoming upset and resistant. They are encouraged to talk about this and how they might have developed such a practice, and then allowed to practise reading together where the parent is warm, positive and rewarding (for a fuller discussion see Pearson and Lindsay, in press). The parent and children are not taught the specific skills of Paired Reading, but the supervised practice allows feedback on the qualities specified above. Relaxed Reading is inherently simpler, and does not require home visiting to monitor the skills learned during Paired Reading. Consequently it could prove to be a more cost-effective form of intervention if it produced improvements comparable to or better than those of Paired Reading.

METHOD

Design

A two-way ANOVA design was used (type of reading intervention × type of monitoring) The Neale Analysis of Reading Ability was used as the dependent measure before and after the interventions.

Subjects

There were 20 children from years M1 and M2 of a middle school in a large northern city. The school contains approximately 200 children aged 8 to 12 years, and serves a mixed area containing owner-occupied, privately rented, and council houses. The sample comprised the 20 children in the M1 and M2 years with the lowest reading ages for Accuracy on the Neale Analysis of Reading Ability. There were 9 boys and 11 girls, 12 children from M1 and 8 from M2 years. The mean age of the sample was 9 years 4 months, range 8 years 7 months to 10 years 4 months.

Procedure

The sample children were allocated to one of four conditions. An attempt was made to achieve a balance in each group between sex and age. Mean levels of reading retardation for the groups ranged from 12.4 to 16.4 months which were not significantly different ($P > 0.05$). The interventions were as follows:

Group A. Paired Reading and Visit Monitoring

These children received a standard paired reading approach. The children and parent were invited to school in the early evening. At least one parent came for each child. The head teacher was present and showed his enthusiasm for the scheme. The parents and children were then introduced to paired reading by one of us (A.E.) who demonstrated by role-play with one of the school's teachers how stressful a parent/child reading session can be. Parents and children were then asked to comment on this, and many remarked that their own experiences had been similar.

Simultaneous Reading was then described and demonstrated by A.E. and the teacher, and questions and comments were invited from parents and children. Then the parent-child pairs practised this using books supplied by the school library. Each pair's Simultaneous Reading was checked by A.E. and when everyone was using the method satisfactorily there was a final discussion. Everyone was keen to continue. They were urged to spend between 10 and 15 minutes per day for 6 or 7 days per week. All parents were then given a Simultaneous Reading reminder sheet and six diary sheets, one for each week of the project. The system of monitoring was explained.

Monitoring was by weekly home visit beginning three days after the first meeting. During these visits the pair was observed reading simultaneously, any faulty technique was corrected and the diary sheets were checked.

A second group meeting was held two weeks after the first when Independent Reading was described and demonstrated as before. Each parent/child pair practised Independent Reading under supervision (A.E.). Each pair was given a reminder sheet for Independent Reading and there was a general discussion.

Group B: Paired Reading and Telephone Monitoring

This group underwent the same training as Group A, at the same time. However, during the six weeks of the intervention home monitoring was by telephone call only. During these calls progress was discussed and a check was made to ensure reading was being done and the diary sheets were being completed.

Group C. Relaxed Reading and Visit Monitoring

This group attended a meeting the evening after the Paired Reading session. The discussion of parental involvement with children's reading was repeated, except that the specific techniques of Paired Reading were not mentioned. The same demonstration of a stressful reading session was presented and resulted in a discussion very similar to that for the Paired Reading groups A and B. Parents and children then practised reading together, under supervision, and the parents were guided to reduce negative and increase positive feedback to the children. No specific methods were taught; rather the parents' normal style of interacting was built on and modified. Monitoring was by weekly home visit starting three days after the first meeting and differed from Paired Reading monitoring only in that the visit checked that Relaxed Reading was being used with no negative comments and that sufficient praise was being given.

Although there was no further technique to be taught a second meeting was held to maintain interest and support, and comprised a general discussion session.

Group D: Relaxed Reading and Telephone Monitoring

This group attended the same meetings as Group C, but as with Group B they received no home visits. Instead discussions were held by telephone to check that Relaxed Reading was being used, and that diary sheets were being completed.

At the end of the six-week period a final meeting was held for the four groups combined.

Measures

The Neale Analysis of Reading Ability was used as a pre- and post-test. On each occasion scores on all three scales were derived: Accuracy, Comprehension and Rate, Parallel Forms (A and B) were used to reduce practice effects, and the order of presentation was randomised across the groups.

RESULTS

(*a*) The progress of the children

(i) *Progress during intervention*

The mean gains or losses in reading age score (in months) on the Neale Analysis of Reading Ability are presented in Table 1. Mean increases in Accuracy range from 5.6 to 10.4 months. For Comprehension the mean increases range from 9.2 to 18 months and the comparable results for Rate are mean losses of 4.0 to 10.6 months. These changes are for a time period of six weeks. Thus the progress of each group as measured by the Accuracy and Comprehension scores of the Neale Analysis of Reading Ability was considerably greater than the 1.5 months increase in chronological age.

TABLE 1

MEAN GAINS IN READING AGE — THE NEALE ANALYSIS OF READING ABILITY (IN MONTHS)

		Accuracy	Comprehension	Rate
Group 1	Paired Reading/Visit	9·0	16·0	−7·2
Group 2	Paired Reading/Phone	10·4	12·0	−4·0
Group 3	Relaxed Reading/Visit	9·0	9·2	−10·6
Group 4	Relaxed Reading/Phone	5·6	18·0	−9·2

(ii) *Progress compared with predicted progress*

These data were analysed to compare the mean gains in reading age for Accuracy and Comprehension with the gains to be predicted on the basis of the children's previous rate of development. The predicted increase in reading age over the 1.5 months was derived from the formula:

$$\frac{RA-6{\cdot}0}{CA-6{\cdot}0} \times 1{\cdot}5$$

where RA = reading age score on the Neale Analysis of Reading Ability; 6.0 represents the minimum reading age level on the Neale; CA = chronological age.

For analysis the Paired Reading groups were combined and the same procedure was carried out for the Relaxed Reading groups. Both interventions were associated with gains in reading age which were significantly greater than those predicted on the basis of the child's reading improvement from the age of six years until the start of the intervention (see Table 2).

TABLE 2

COMPARISON OF THE MEAN SCORES OF ACTUAL RATE OF READING PROGRESS PER MONTH *VS* PREDICTED RATE OF PROGRESS

		Actual gain	Predicted gain	t	P
Paired Reading	Accuracy	6·47	0 97	63 73	0 0000
Paired Reading	Comprehension	8 10	0·84	51·22	0·0000
Relaxed Reading	Accuracy	4 67	1 01	51 73	0·0000
Relaxed Reading	Comprehension	8 17	0·98	73·35	0·0000

(iii) *Comparison of Paired Reading and Relaxed Reading*

A two-way ANOVA was performed on the data derived from the four conditions for each of the three reading scores as dependent variables. In all cases the F ratios failed to achieve significance (see Table 3).

(iv) *Rate of reading*

The mean reading ages for rate of reading all show significant decreases, ranging from 4·0 to 10·6 months.

TABLE 3

EFFECTS OF INTERVENTIONS ON READING

Reading Measure	Intervention	F ratio
Accuracy	Paired *vs.* Relaxed Reading	1 67
	Phone *vs.* Home Visiting	0 29
	Interaction	1·67
Comprehension	Paired *vs.* Relaxed Reading	0·01
	Phone *vs.* Home Visiting	0 25
	Interaction	1·80
Rate	Paired *vs.* Relaxed Reading	0 94
	Phone *vs* Home Visiting	0·37
	Interaction	0

N.B. — All F ratios are non-significant, df = 1 in all cases.

DISCUSSION

Earlier studies have suggested that Paired Reading can produce large gains in reading performance over a short period of intervention. This study confirms these findings. However, the study goes further and demonstrates that a much simpler type of intervention, Relaxed

Reading, can also produce significant improvements. For both conditions the gains made during the six weeks intervention were significantly greater than would be expected on the basis of an increase in chronological age. Of more interest is the comparison with the children's previous rates of improvement in reading performance, which were very low. For example, the mean increase in Accuracy of 5·57 months for the total sample compares with the 0·99 months to be expected by the children's previous performance.

However, the study also suggests that a simpler form of intervention, Relaxed Reading, was as effective in improving the children's reading ability, as no significant differences were found in main effects. This has important practical implications as Relaxed Reading is simpler to teach to parents and does not require the time-consuming home visiting usually accompanying Paired Reading.

The greater improvement in reading found following either Relaxed or Paired Reading compared with that following simple parental involvement schemes lends support to the importance of the structured approach to the project. This includes specific meetings, discussion of approaches and monitoring as outlined above. However, the relative success of Relaxed as opposed to Paired Reading suggests that the specific teaching approach used in the latter is not crucial.

The decrease in the mean reading age for Rate appears, at first sight, to be an anomalous result. Mean scores for accuracy and comprehension have both increased while mean scores for rate have fallen. The present study does not allow a detailed examination of this phenomenon. What is of interest, however, is that mean increases in accuracy did not impair increases in comprehension; improvement in accuracy developed in parallel with that for comprehension. It might be speculated that this is a result of the children slowing down their reading (as shown by the reduction in mean Rate scores) which in turn aided accuracy directly, and comprehension either or both directly, or indirectly via the improvement in accuracy. Further studies would be necessary to investigate possible mechanisms, if this effect is replicable.

These results must be tempered with some caution as the sample size is small and adequacy of the measures is open to dispute. A reading age measured in months is a crude instrument for an intervention over six weeks. However, the findings replicate earlier work which has examined Paired Reading, and found similar significant main effects. Of more concern is the technical adequacy of the Comprehension Scale of the Neale where a child may achieve a large increase in reading age owing to a test artefact (see Pearson and Lindsay, in press).

A final technical point concerns the use of gain scores. Those reported here are simple gains in reading age. A further set of analyses was executed on gain scores controlled for initial reading age. These results were very similar to those reported here.

Thus, this study, despite some caveats, supports the body of work which has demonstrated the powerful effects of short-term intervention by parents in improving their child's reading. The evidence on Paired Reading is now very substantial with many studies reporting significant gains. The success of Relaxed Reading suggests that this less formal approach, which has much in common with Paired Reading, can produce similar benefits. Relaxed Reading, therefore, would be the preferred approach as it is simpler to introduce and easier to implement.

REFERENCES

BUSHELL, R., MILLER, A., and ROBSON, D. (1982) Parents as remedial teachers *Assoc. Educ. Psychologists J.*, 5 (9), 7-13.

HANNON, P. W., JACKSON, A., and PAGE, B. (1985). Implementation and take-up of a project to involve parents in the teaching of reading. In TOPPING, K., and WOLFENDALE, S. (Eds), *Parental Involvement in Children's Reading*. London: Croom Helm.

HEWISON, J., and TIZARD, J. (1980). Parental involvement and reading attainment. *Br. J. educ. Psychol.*, 50, 209-215.

JACKSON, A., and HANNON, P. W. (1981). *The Belfield Reading Project*. Rochdale: Belfield Community Council.

MORGAN, R., and LYON, E. (1979). Paired Reading — a preliminary report on a technique for parental tuition of reading retarded child. *J child Psychol. Psychiat*, 20, 151-160.

NEALE, M. (1957) *The Neale Analysis of Reading Ability* Basingstoke: Macmillan.

PEARSON, L., and LINDSAY, G (in press). *Special Needs in the Primary School: Identification and Intervention* Windsor· NFER-Nelson.

TIZARD, J., SCHOFIELD, W. N., and HEWISON, J. (1982). Collaboration between teachers and parents in assisting children's reading. *Br J. educ. Psychol.*, 52, 1-15.

TOPPING, K., and McKNIGHT, G. (1984) Paired Reading — and parent power. *Special Education: Forward Trends*, 11 (3), 12-14.

(*Manuscript received June*, 1985)

Br. J. educ. Psychol., 55, 310-313, 1985

THE USE OF LONGITUDINAL DATA IN A STUDY OF CHILDREN'S SCHOOL MOBILITY AND ATTAINMENT

BY D. C. BLANE
(Faculty of Education, Monash University, Victoria, Australia)

DORIA PILLING AND K. FOGELMAN
(National Children's Bureau, London)

SUMMARY. Longitudinal data contained in the National Child Development Study (NCDS) cohort was used in a Department of Education and Science (DES) funded investigation of whether changes of school adversely affect children's attainment and behaviour. There have been a number of studies which have investigated these effects but the results have been conflicting and confusing Recently a series of major studies has shown that it may be especially dangerous to draw conclusions based on *ex post facto* data in this field. This research note discusses some aspects of the research methodology adopted in this study to overcome the deficiences of some of the earlier mobility projects The findings, based on a series of multivariate analyses, support the view that differences in the academic attainment of children who change school generally existed prior to the move.

INTRODUCTION

Geographic mobility has always been a significant factor in the lives of many differing occupational and other groups and has been perceived by a large number of parents and teachers as a major difficulty, particularly when school changes occur. Children from military families, from single parent families, those with unemployed fathers as well as successful fathers in professional and managerial occupations are all groups for whom frequent changes of school can be relatively common and who may be at risk from the effects of regular changes of home and school Reviews of the literature and previous research, most recently by Blane (1983), reveal a confusing picture of conflicting findings and suggestions for combating the detrimental educational and other effects thought to arise from child mobility, especially in relation to schooling and academic achievement. Despite the lack of agreement of the research findings in this area there have been quite definitive statements in several authoritative publications which were highlighted by Lacey and Blane (1979). These authors, however, concluded from the evidence available that the effects of school mobility could only be adequately assessed using longitudinal data.

Earlier Schaller (1972) had suggested that research on mobility should concentrate on the three phases of mobility: the pre-mobile situation of the child, factors intimately related to moving and moving decisions, and the child's situation after moving. In a subsequent study, Schaller (1976) tested his hypothesis that differences in academic performance between children who had never moved and children who had moved several times before to a great extent stem from differences already existing prior to mobility. Schaller's results served as a warning against drawing hasty conclusions from research based on *ex post facto* data and were influential in the recommendations by Lacey and Blane (1979) and Blane (1980) which resulted in the decision by the Department of Education and Science (DES) in 1982 to fund a major study in this area using the data contained in the National Child Development Study (NCDS). The detailed results of this DES investigation of school mobility and attainment are contained in a report by the National Children's Bureau (1983).

METHOD

Sample

The data on which this investigation was based were drawn from those held as part of the NCDS at the National Children's Bureau. The NCDS is a long term study carried out by the National Children's Bureau of all the young people born in the week 3rd-9th March, 1958 and living in England, Scotland and Wales. The study, its methods of data collection and the variables are described fully in Fogelman (1983). The variables used in the following analyses are taken from information obtained about the children when they were aged 7, 11 and 16, mainly from parental interview but also from the child's school, together with reading comprehension and mathematics tests administered for the study by schools at the same time.

The mobility variable is based on a question used in the parental interview at 16 which asked the parent to name all the schools attended since the age of 11. Only changes of school beyond those taking place as a normal part of the educational system were counted. This excluded transfers at 12 or 13 if these were the normal procedure, as in Scotland, or between a preparatory and independent secondary school. Children who passed through a middle school system were excluded altogether from these analyses. Table 1 shows the distribution of children who attended various numbers of secondary schools who are contained in the cohort.

TABLE 1

NUMBER OF SCHOOLS ATTENDED IN SECONDARY (11-16) PERIOD

	N	%
One school	6863	83·0
Two	1202	14·5
Three	173	2·1
Four or more	32	0·4

Procedure

The first aim of this study was to provide descriptive comparisons of children who had and had not changed school and two separate investigations were made during their primary and secondary school periods. The full descriptive results can be obtained from the Report together with details of the various statistical analyses carried out on a range of variables not included in these notes. The main purpose of the descriptive section of the investigation was, however, to provide some guide to the differences in background factors which would need to be allowed for in multivariate analyses which were attempting to isolate the independent effect of school mobility on various measures of attainment and behaviour.

The descriptive analyses confirmed the need to carry out the further investigation with multivariate techniques, if the effects of school mobility were to be isolated. Not unexpectedly, the descriptive analyses failed to provide the clear-cut answer to the question of whether mobility does adversely affect attainment. For the most part the attainment scores of the more mobile children were lower than those of the non-mobile children, but usually these differences pre-existed at the start of the period of schooling under consideration, before the changes of school had occurred, and seemed to confirm Schaller's (1976) findings. The secondary school data shown in Table 2 indicate that the mathematics test scores decreased with mobility, and that there were significant differences between the mobility groups. There was a similar pattern for reading comprehension, although the differences between the mobility groups were not statistically significant.

TABLE 2

ATTAINMENT AT 16 BY NUMBER OF SECONDARY SCHOOLS ATTENDED

	Reading Comprehension		Mathematics	
	Mean	SD	Mean	SD
One school	25·8	6·74	13·2	6·96
Two	25·4	7·22	13·0	7·21
Three	24·6	7·13	11·9	6·43
Four or more	26·0	6·91	10·3	7·13

A similar pattern can be seen in Table 3, however, for these same children at age 11, suggesting that to some extent the differences pre-existed.

The purpose of the multivariate analyses was to adjust for initial differences in attainment and social or home circumstances between mobile and non-mobile children at the beginning of the period of schooling under consideration and to see whether there were then any remaining differences between those attending different numbers of schools during this period on various measures of attainment and behaviour.

TABLE 3

ATTAINMENT AT 11 BY NUMBER OF SECONDARY SCHOOLS ATTENDED

	Reading Comprehension		Mathematics	
	Mean	SD	Mean	SD
One school	16·3	6·23	17·4	10·23
Two	16·0	6·46	16·9	10·44
Three	15·2	6·47	16·3	10 29
Four or more	17 2	6 24	15·0	9·81

Table 4 summarises the results of multiple regression with the independent variables listed under the table, and shows the 'effect' of mobility when this is entered as 'last one in'. In addition to the children attending middle schools, the following groups of children were excluded from the multivariate analyses:

(i) Children who attended four or more secondary schools. This group was too small for separate analysis, and, as these children were found to differ in a number of characteristics from other mobile children, it was decided against combining them with those attending three schools. The full report contains a summary of a case history study of this group.
(ii) Children who, at the age of 11, were attending special schools.
(iii) Children with fathers in the armed forces, who could not therefore be assigned to a social class.

TABLE 4

SECONDARY SCHOOL MOBILITY AND 16-YEAR ATTAINMENT — ANALYSIS OF VARIANCE

	Reading Comprehension			Mathematics		
Total Variance	0 898			0·947		
N	5941			5924		
Residual Mean Square	0·315			0·406		
	Fitted constants	df	X^2	Fitted constants	df	X^2
One school	0·018			0·043		
Two schools	−0·029	2	4·6	−0·001	2	4·2
			(P>0 05)			(P>0·05)
Three schools	0·011			−0·042		

NOTE· Other independent variables in the above analyses: Social class at 11; sex; whether school attended at 16 was maintained or not; tenure of family's home at 11; whether any child in the family received free school meals at 11; reading or mathematics test score at 11.

In interpreting this table the differences between fitted constants can be taken as equivalent to the differences between adjusted means. The χ^2 distribution is identical with the F-test for this large number of degrees of freedom. To satisfy the assumptions of this kind of analysis the 16-year test scores have been transferred to a scale with a mean of zero and an SD of one.

DISCUSSION

Overall the findings of the multivariate analyses confirm the results of other research in so far as this has found the effects of school mobility on children's attainment to be only small or marginal. It can be seen, therefore, that differences between different mobility groups are largely explained by the pre-existing differences in attainment together with differences in other characteristics and there is little evidence of an increase in adverse effects with mobility at the secondary school level.

These results and those contained in the main report further strengthen the warning offered by Schaller (1976) against drawing hasty conclusions from studies of mobility and academic performance based only on *ex post facto* data. They further support the hypothesis that the differences in academic performance after a move, between those children who do and do not change school, depend to a large extent on differences in their performance prior to the move.

REFERENCES

Blane, D C (1980). A report to the DES Boarding Education Group reviewing the available research evidence on the relation between children's geographic mobility and school attainment. In *Boarding Education*. London: Department of Education and Science.

Blane, D. C (1983). A review of the research on geographic mobility and its effect on children. In *School Mobility and Attainment*. London National Children's Bureau

Fogelman, K. (Ed.), (1983). *Growing Up in Great Britain*. London· Macmillan.

Lacey, C. and Blane, D. C (1979). Geographic mobility and school attainment — the confounding variables *Educ. Res.*, 21, 200-206.

National Children's Bureau (1983). *School Mobility and Attainment*. London: National Children's Bureau.

Schaller, J., (1972). Residential change and various factors of school adjustment· a review of research. *Göteborg Psychol Rep.*, 2, 21.

Schaller, J., (1976). Geographic mobility as a variable in ex-post facto research. *Br J educ. Psychol.*, 46, 341-343.

(*Manuscript received* 15*th March*, 1984)

CALCUTTA CENTRAL LIBRARY

PUBLICATIONS RECEIVED

Books for review have been selected from the following
('n.p.' means 'no price mentioned').

Adams, F. (Ed.) (1985). *Science and Computers in Primary Education.* Edinburgh: Scottish Council for Research in Education, pp. 173, £9·00.

Ainley, J., Batten, M., and Miller, H. (1984). *Patterns of Retention in Australian Government Schools. ACER Research Monograph No. 27.* Victoria: Australian Council for Educational Research, pp. 135, $10·00.

Ashman, A. F., and Laura, R. F. (Eds.) (1985). *The Education and Training of the Mentally Retarded. Recent Advances.* London: Croom Helm, pp. 295, £14·95 hb.

Barrett, M. D. (Ed.) (1985). *Children's Single-Word Speech.* Chichester: John Wiley and Sons, pp. 311, £28·00.

Bee, H. (1985). *The Developing Child.* (4th edit.) London: Harper and Row, pp. 582, £10·50.

Bennett, N., and Desforges, C. (Eds.) (1985). *Recent Advances in Classroom Research.* Edinburgh: Scottish Academic Press, pp. 201 (Monograph Series No. 2), n.p.

Berkson, W., and Wettersten, J. (1984). *Learning from Error: Karl Popper's Psychology of Learning* (translation). Illinois: Open Court Publishing Company, pp. 155, $14·95.

Berridge, D. (1985). *Children's Homes.* Oxford: Basil Blackwell, pp. 128, £16·50 hb, £7·50 pb.

Black, H. D., Alexander, W., Boyle, E., and Goring, R. (1984). *Administering Change in Assessment. A Local Authority Case Study.* Edinburgh: Scottish Council for Research in Education, pp. 103, n.p.

Black, H. D., and Dockrell, W. B. (1984). *Criterion Referenced Assessment in the Classroom.* Edinburgh: Scottish Council for Research in Education, pp. 225, £12·50 hb, £8·00 pb.

Bourke, S. F. (1984). *The Teaching and Learning of Mathematics.* Victoria: Australian Council for Educational Research, pp. 256, $15·00.

Bradley, L., and Bryant, P. (1985). *Rhyme and Reason in Reading and Spelling.* Michigan. University of Michigan Press, pp. 120, $7·95 pb.

Branthwaite, A., and Rogers, D. (1985). *Children Growing Up.* Milton Keynes: Open University Press, pp. 246, £5·95 pb.

Bruner, J. (1984). *In Search of Mind. Essays in Autobiography.* London: Harper and Row, pp. 293, $6·95.

Case, R. (1985). *Intellectual Development: Birth to Adulthood.* London: Academic Press, pp. 429, £39·50.

Collins, A. W. (1984). *Development During Middle Childhood: The Years From Six To Twelve.* Washington DC: National Academy Press, pp. 419, £35·90.

Cooper, B. (1985). *Renegotiating Secondary School Mathematics.* Basingstoke: Falmer Press, pp. 283, £15·95 hb, £8·95 pb.

Corsini, R. J. (Ed.) (1984). *Encyclopedia of Psychology* (4 vols). Chichester: Wiley, £189·95.

Davis, L. (1985). *Caring for Secondary School Pupils.* London: Heinemann Educational, pp. 101, n.p.

Dowling, E., and Osborne, E. (Eds.) (1985). *The Family and The School A Joint Systems Approach to Problems with Children.* London: Routledge and Kegan Paul, pp. 203, £7·95.

Duane, D. D., and Leong, C. K. (Eds.) (1985). *Understanding Learning Disabilities: International and Multidisciplinary Views.* New York: Plenum, pp. 269, $42·50.

Duffy, T. M., and Waller, R. (Eds.) (1985). *Designing Usable Texts.* London: Academic Press, pp. 400, £48·00.

Eichenbaum, L., and Orbach, S. (1985). *Understanding Women.* London: Penguin Books, pp. 212, £2·95.

Elliott, M. (1985). *Preventing Child Sexual Assault: a Practical Guide to Talking with Children.* London: Bedford Square Press/NCVO, pp. 44, £1·95.

Ericksen, S. C. (1985). *The Essence of Good Teaching.* London: Jossey-Bass, pp. 180, £22·00.

Eysenck, H. J., and Eysenck, M. W. (1985). *Personality and Individual Differences: A Natural Science Approach.* New York: Plenum, pp. 358, £29·50

Farrish, S. J. (1984). *Investigating Item Stability.* Victoria: Australian Council for Educational Research, pp. 96, $9·00.

Fish, J. (1985). *Special Education: The Way Ahead.* Milton Keynes: Open University Press, pp. 145, n.p.

Fontana, D. (Ed.) (1984). *Behaviourism and Learning Theory in Education.* Edinburgh: Scottish Academic Press, pp. 195 (Monograph Series No. 2), n.p.

Freeman, J. (Ed.) (1985). *The Psychology of Gifted Children: Perspectives on Development and Education.* Chichester: John Wiley and Sons, pp. 394, £24·75.

Freud, S. (1985). *13. The Origins of Religion.* London: Penguin Books, pp. 413, £4·95.

Fry, P. S. (Ed.) (1984). *Changing Conceptions of Intelligence and Intellectual Functioning: Current Theory and Research.* Amsterdam: Elsevier, pp. 226, $44·25.

Galloway, D. (1985). *Schools, Pupils and Special Educational Needs.* London: Croom Helm, pp. 186, £6·95 pb.

Grzesik, J., and Fischer, M. (1984). *Was leisten Kriterien fur die Aufsatzbeurteilung?* Wiesbaden 2: Friedr. Vieweg and Sohn, pp. 253, DM 41.

Harris, R. J., Barker, M. W., Reading, P., Richards, M., and Youll, P. (Eds.) (1985). *Educating Social Workers.* Leicester: Association of Teachers in Social Work Education, pp. 147, £7·95.

Harth, E. (1985). *Windows on the Mind. Reflections on the Physical Basis of Consciousness.* Harmondsworth: Penguin Books, pp. 285, £4·95.

Hill, W. F. (1985). *Learning: A Survey of Psychological Interpretations* (4th edit.). London: Harper and Row, pp. 220, £14·95 pb.

Jung, C. G., and Kerenyi, C. (1985). *Science of Mythology: Essays on the Myth of the Divine Child and the Mysteries of Eleusis.* London: Ark, pp. 200, £3·95.

Kahle, J. B. (1985). *Women in Science: a Report from the Field.* Lewes: Falmer Press, pp. 251, £7·95 pb.

Katz, L. G. (1984). *Current Topics in Early Childhood Education, Vol. V.* London: Ablex, pp 234, n.p.

Larkin, A. I., and Keeves, J. P. (1984). *The Class Size Question: A Study at Different Levels of Analysis.* Victoria: Australian Council for Educational Research, pp. 119, n.p.

Laura, R. S., and Ashman, A. F. (Eds.) (1985). *Moral Issues in Mental Retardation.* London: Croom Helm, pp. 212, £14·95 hb.

Lavender, G. B., and Findlay, M. A. (1985). *Australian Thesaurus of Education Descriptors.* Victoria: Australian Council for Educational Research, pp. 268, $39·00.

Lawler, R. W. (1985). *Computer Experience and Cognitive Development: A Child's Learning in a Computer Culture.* Chichester: Ellis Horwood, pp. 247, £25·50.

MacKinnon, G. E., and Waller, T. G. (Eds.) (1985). *Reading Research: Advances in Theory and Practice, Vol. 4.* London: Academic Press, pp. 262, £35·00.

Marklund, I. (1984). *Formulating a National Educational Research Programme: The Case of Sweden.* Edinburgh: Scottish Council for Research in Education, pp. 39, n.p.

Marton, F., Hounsell, D., and Entwistle, N. (Eds.) (1984). *The Experience of Learning.* Edinburgh: Scottish Academic Press, pp. 242, n.p.

McConkey, R. (1985). *Working with Parents A Practical Guide for Teachers and Therapists.* Beckenham: Croom Helm, pp. 304, £19 95 hb, £9·95 pb.

McTear, M. (1985). *Children's Conversation.* Oxford: Basil Blackwell, pp. 295, n.p.

Meek, V. L. (1984). *A Study of the Gippsland Institute of Advanced Education: Brown Coal or Plato?* Australia: Australian Council for Educational Research, pp. 269, $15·00.

Montgomery, D. (1984). *Evaluation and Enhancement of Teaching Performance.* London: Kingston Polytechnic, pp. 107, n.p.

Murgatroyd, S., and Woolfe, R. (1985). *Helping Families in Distress. An Introduction to Family Focussed Helping.* London: Harper and Row, pp. 154, £5·95 pb.

Neill, A. S. (1985). *Summerhill.* London. Penguin Books, pp. 336, £3·95.

Nisbet, J., and Shucksmith, J. (1984). *The Seventh Sense: Reflections on Learning to Learn.* Edinburgh: Scottish Council for Research in Education, pp. 16, £1 60.

Norris, D. E., Skilbeck, C. E., Hayward, A. E., and Torpy, D. M (1985). *Microcomputers in Clinical Practice.* Chichester: Wiley, pp. 159, £7·90.

Peacock, A., and Denvir, K. (1985). *Scottish Postgraduate Research into Maladjustment 1974-1983.* Edinburgh: Scottish Council for Research in Education, pp. 23, £1·50.

Perfetti, C. A. (1985). *Reading Ability.* Oxford: Oxford University Press, pp. 282, £25·00.

Pichot, P., Berner, P., Wolf, R., and Thau, K (Eds.) (1985). *Psychiatry: The State of the Art.* New York: Plenum, pp. 498, $55 00 hb.

Pitcher, E. G., and Schultz, L. H. (1983). *Boys and Girls at Play: the Development of Sex Roles.* London: Bergin and Garvey, pp. 176, £27 95 cloth, £12·95 pb.

Pollitt, A., Entwistle, N., Hutchinson, C., and De Luca, C. (1985). *What Makes Exam Questions Difficult. An Analysis of 'O' Grade Questions and Answers.* Edinburgh: Scottish Academic Press, pp. 185, £8·75.

Powell, J. L. (1985). *The Teacher's Craft: a Study of Teaching in the Primary School.* Edinburgh: Scottish Council for Research in Education, pp. 258, £13·50 hb, £8·95 pb.

Powell, J. L. (1985). *Ways of Teaching.* Edinburgh: Scottish Council for Research in Education, pp. 38, £2·00.

Powell, T. H., and Ogle, P. A. (1985). *Brothers and Sisters: A Special Part of Exceptional Families.* London: Brookes, pp. 226, n.p.

Raven, J., with Johnstone, J. and Varley, T. (1985). *Opening the Primary Classroom.* Edinburgh: Scottish Council for Research in Education, pp. 151, £12·50 hb, £8·00 pb.

Rivlin, L. G., and Wolfe, M. (1985). *Institutional Settings in Children's Lives.* Chichester: John Wiley & Sons, pp. 231, £31·00.

Rodger, I. A., and Richardson, J. A. S. (1985). *Self-evaluation for Primary Schools.* Sevenoaks: Hodder and Stoughton, pp. 198, £4·95 pb.

Ross-Jones, F., and Morgan, R. F. (1985). *The Psychology of Human Development.* London: Harper and Row, pp. 499, £31·25 pb.

Rowe, H. (1985). *Problem Solving and Intelligence.* London: Lawrence Erlbaum Associates, pp. 345, £36·00.

Ryrie, A. C. (1984). *The New Approach in Further Education.* Edinburgh: Scottish Council for Research in Education, pp. 74, n.p.

Schopler, E., and Mesibov, G. B. (Eds.) (1985) *Communication Problems in Autism.* New York: Plenum, pp. 325, $35·00.

Schwartz, S. (1984). *Measuring Reading Competence.* New York: Plenum, pp. 337, $29·50.

Segal, J. (1985). *Phantasy in Everyday Life.* Harmondsworth: Penguin, pp. 234, £2·95.

Shulman, V. L., Restaino-Baumann, L. C. R., and Butler, L. (Eds.) (1985). *The Future of Piagetian Theory: The Neo-Piagetians.* New York: Plenum, pp. 216, $29·50 hb.

Simon, B. (1985). *Does Education Matter?* London: Lawrence and Wishart, pp. 228, £5·95 pb.

Slavin, R., Sharan, S., Kagain, S., Lazarowitz, R. H., Webb, C., and Schumuck, R. (1985). *Learning to Co-operate, Co-operating to Learn*, pp. 472, $45·00.

Thomas, L. F., and Harri-Augstein, E. S. (1985). *Self Organised Learning: Foundations of a Conversational Science for Psychology.* London: Routledge and Kegan Paul, pp. 388, £25·00.

Topping, K., and Wolfendale, S. (Eds.) (1985). *Parental Involvement in Children's Reading.* London: Croom Helm, pp. 333, £12·95 pb.

Webster, A., and Ellwood, J. (1985). *The Hearing Impaired Child in the Ordinary School.* Beckenham: Croom Helm, pp. 181, £14·95 hb, £8·95 pb.

Wilson, J. A. (1985). *Secondary School Organisation and Pupil Progress.* Belfast: NICER, pp. 95, n.p.

Winn, M. (1985). *The Plug-In Drug: Television, Children and the Family.* London: Penguin, pp. 288, £3·95.

AUTHOR INDEX TO VOLUME 55, 1985

ABRAMS, D : Gender Salience and Social Identity. the Impact of Sex of Siblings on Educational and Occupational Aspirations ... 224

BANKS, M H . see Johnson, M. R. D.

BEECH, J. R : see Harding, L. M.

BEVERIDGE, M.: The Development of Young Children's Understanding of the Process of Evaporation ... 84

BEWLEY, B. R · see Johnson, M R. D.

BIGGS, J B.· The Role of Metalearning in Study Processes ... 185

BIGSBY, P.: The Nature of Reversible Letter Confusions in Dyslexic and Normal Readers: Misperception or Mislabelling? ... 264

BLAND, J M : see Johnson, M. R. D

BLANE, D. C.: The Use of Longitudinal Data in a Study of Children's School Mobility and Attainment ... 310

BOREHAM, N. C.: Transfer of Training in the Generation of Diagnostic Hypotheses· the Effect of Lowering Fidelity of Simulation ... 213

BORG, M.: see Wheldall, K.

CLYDE, D. V.: see Johnson, M. R. D.

COOPER, C see Kline, P.

CRASKE, M. L.. Improving Persistence through Observational Learning and Attribution Retraining 138

DARCH, C.: The Effects of Teacher Presentation Rate and Praise on LD Students' Oral Reading Performance ... 295

ENTWISTLE, N J : Relationships between School Motivation, Approaches to Studying, and Attainment among British and Hungarian Adolescents ... 124

EVANS, A.: see Lindsay, G.

FOGELMAN, K · see Blane, D. C.

FREEBODY, P : Achievement Outcomes of Two Reading Programmes: An Instance of Aptitude-Treatment Interaction ... 53

GERSTEN, R.. see Darch, C.

GIBBS, J.: The Validity of a Short Form of the Coopersmith Self Esteem Inventory. ... 76

GILL, C. E.: Further Evidence for Genetic Influences on Educational Achievement ... 240

HALL, E.: see Woodhouse, D. A.

HANNELL, G.: see Yates, G. C. R.

HARDING, L. M : The Changing Pattern of Reading Errors and Reading Style from 5 to 11 Years of Age ... 45

HART, J. G : LAWSEQ: Its Relation to Other Measures of Self-Esteem and Academic Ability .. 167

HODGSON, S : see Rusted, J.

HOFFMAN, M. A.: see Levy-Shiff, R.

HOGG, M A.. see Abrams, D.

JARDINE, R.: see Gill, C. E.

JOHNSON, M. R. D.: Schools and Smoking: School Features and Variations in Cigarette Smoking by Children and Teachers ... 34

JONES, B.: see Lindsay, G.

KEMBER, P.: The Transcendental Meditation Technique and Post-Graduate Academic Performance 164

KLINE, P.: Rigid Personality and Rigid Thinking ... 24

KOZÉKI, B.: see Entwistle, N. J.

KRAFT, G. S.: Verbal/Spatial Task Expectations in Written Instruction ... 273

KULHAVY, R. W.: see Kraft, G. S.

Kyriacou, C.: Conceptualising Research on Effective Teaching ... 148

Kyriacou, C. Teacher Stress and Psychoneurotic Symptoms ... 61

Levy-Shiff, R.: Social Behaviour of Hearing-Impaired and Normally-Hearing Preschoolers ... 111

Lindsay, G.: Paired Reading versus Relaxed Reading: A Comparison ... 304

Lippett, R. M.: see Yates, G. C. R

Little, A. W.: The Child's Understanding of the Causes of Academic Success and Failure: A Case Study of British Schoolchildren ... 11

Low Guat Tin: An Observational Study of the Social Adjustment of Spina Bifida Children in Integrated Settings ... 81

McKay, M. F.: The Predictive Validity of Bannatyne's WISC Categories for Later Reading Achievement ... 280

Martin, N. G see Gill, C. E.

Merrett, F.. see Wheldall, K.

Mevarech, Z. R.. The Relationships between Temperament Characteristics, Intelligence, Task-Engagement and Mathematics Achievement ... 156

Neale, M. D : see McKay, M. F.

Norwich, B.: see Gibbs, J.

Novie, G. J.: see Kraft, G. S.

Pilling, D.: see Blane, D. C.

Pratt, J.: see Kyriacou, C. (second entry)

Richmond, P. G.: The Relationship of Grade, Sex, Ability and Socio-Economic Status to Parent, Peer and School Affiliation ... 233

Rusted, J.: Evaluating the Picture Facilitation Effect in Children's Recall of Written Texts ... 288

Sneddon, W : see Harding, L M.

Sparkes, K see Abrams, D

Szaday, C. S B. see Youngman, M B.

Teasdale, G R see Low Guat Tin

Thomas, D.: The Determinants of Teachers' Attitudes to Integrating the Intellectually Handicapped 251

Thompson, G B.. see McKay, M. F.

Tirre, W C.. see Freebody, P.

Wheldall, K: The Behavioural Approach to Teaching Package (BATPACK): an Experimental Evaluation ... 65

Woodhouse, D. A: Taking Control of Stress in Teaching ... 119

Wooster, A D : see Woodhouse, D A.

Yates, G C. R. Cognitive Slippage, Test Anxiety and Responses in a Group Testing Situation ... 28

Youngman, M B.: Further Validation of the Behaviour in School Inventory ... 91

Zakay, D The Influence of Information and Daily Contact on Children's Attitudes Towards Aphasic Children ... 1

NOTES TO CONTRIBUTORS

Manuscripts are accepted on the understanding that they report unpublished work that is not under consideration for publication elsewhere, and that if accepted for the *Journal*, the work will not be published again in the same form, in any language, without the consent of the Editor.

The usual range of length of papers is from 2,500 to 4,000 words. Papers should be in a clear style and free from jargon Statistical matter (formulae, tables, discussion of techniques, etc.) should be kept to a minimum, so that the *Journal* will be of interest to statistically untrained readers. Such concepts as mean, percentiles, standard deviations, correlation, chi-squared, standard error, critical ratio, significance and reliability, regression and factor analysis, can be employed in the text and need not be explained.

Short research notes will also be published, reporting work of two kinds: (i) results of sufficient moment to merit publication in advance of a more comprehensive paper; (ii) work which substantially' confirms or extends existing knowledge but which does not justify an extensive paper.

Contributors can help to avoid delays in publication by observing certain guiding principles.

1—Papers should be written with the utmost conciseness consistent with clarity. If the editor or a referee takes the wrong meaning from or fails to understand a passage, the fault is the author's. Unless the editor can grasp the meaning of a sentence unequivocally, it cannot be assumed that other readers can.

2—A paper should be written only when a piece of work is rounded off A report of a pilot study, or a series of papers on the same subject written as the results come to hand, is unlikely to be accepted for publication. A comprehensive paper, particularly if it includes reports of several experiments or analyses, or replications on further samples, which give consistent results, will be given favourable consideration.

3—Experimental reports should generally follow the standard pattern of presentation, e.g .

(*a*) A short summary at the start.
(*b*) A statement of aims or objectives and a review of relevant previous work as an introduction to the topic—conciseness and strict selection of relevant aspects are particularly important here.
(*c*) A clear statement of the sample, the procedure, the tests, the experimental material, in sufficient detail to permit a reader to replicate the study if he so wishes
(*d*) A factual statement of results, without speculative comment or discussion.
(*e*) Either a brief list of general conclusions, not repeating results already listed, or a discussion, which should be strictly limited to discussing the results (if this is necessary) and which should not involve mere recapitulation or speculation beyond the scope of the inquiry.
(*f*) Acknowledgments.
(*g*) References.

4—The summary (which is often copied by abstracting agencies) should be intelligible without reference to the paper. It should give an indication of the aim, scope and results of the inquiry, preferably in about three sentences. It should mention the size and nature of the sample, and should not include vague phrases such as "the results were analysed . ." or "The implications of the findings are discussed . ." Even for a full-length paper, the summary should not exceed 200 words.

5—Figures and tables should be selected to illustrate or summarise points made in the text The same data should not be reported both in table and figure; tables are generally preferable because they are much more economical Tables should contain selected data. only the most important quantitative results can be reported in full

6—Authors should consult a current issue of the *Journal* in order to follow the conventions applied in the *Journal*—the use of headings, numbering of tables, the absence of footnotes, the impersonal style of reporting, omission of dots in standard abbreviations, reference to subjects as children and not as Ss, the lay-out of references and so on The British Psychological Society pamphlet, *Suggestions to Contributors*, should be consulted. References should be cited by author and date in the text, and arranged in alphabetical order in the list of references at the end. All measurements should be stated in SI units. The *Journal* is committed to the use of non-sexist language.

7—A manuscript should be headed with the title of the paper, the names of the authors and, in brackets, a short address identifying the institution where the work was done Titles should be short. If the title exceeds 45 letters and spaces in length, a shortened version, suitable as a running title, should be added.

8—The manuscript should be typed in double-spacing with wide margins on one side of the paper which is suitable for ink corrections and printer's instructions—not duplicated or on flimsy paper. Drawings should be in Indian ink on cartridge paper Manuscripts which require re-typing before they can be handled by the printer have to be sent back to the authors. The *Journal* has no paid staff.

9—Authors, if they specifically request, may have their papers reviewed blind.

Authors receive 50 copies of their papers free (delivered about six weeks after publication) Extra copies may be had at cost price if the order is given when the proof is returned.

The *Journal* will adopt the International System of Units (SI) based on the units. metre, kilogramme, second, ampere, Kelvin and candela Authors who refer to physical measurements in their papers should now use SI units except where it would be absurd to do so Further information about SI units, including conversion tables, is contained in the revised *Suggestions to Contributors* pamphlet issued by the British Psychological Society and obtainable, price £2.00 post free, from The British Psychological Society, St Andrew's House, 48 Princess Road East, Leicester LE1 7DR.

THE BRITISH JOURNAL OF EDUCATIONAL PSYCHOLOGY
MONOGRAPH SERIES

BEHAVIOURISM AND LEARNING THEORY IN EDUCATION

David Fontana

Education is notoriously eclectic and practitioners have consistently used both behaviouristic and cognitive notions to help their understanding of children in schools. The authors of *Behaviourism and Learning Theory in Education* suggest it is time for educational psychologists to follow suit, for in a sense all psychologists and all educationalists are behaviourists in that they study how people actually behave and how this behaviour can be changed by the action of others.

This Monograph surveys the contribution currently made by behaviourism and learning theory within formal education and within those fields, such as counselling, closely associated with it. It is not intended as an apologia for behaviourism *per se*, but rather as a demonstration that behaviourism is a useful tool within education, when interpreted and applied correctly.

The text is divided into two sections. Section I deals with behaviourism and learning theory in the classroom with particular emphasis on the day-to-day contacts between teacher and child. Section II looks at broader issues, such as language disabilities, which engage the educationalist on a less routine (but equally important) basis. Each section is supplemented with an editorial which presents a critical overview of the preceding chapters, and the Monograph concludes with a discussion on the future of behaviourism and learning theory.

SBN 7073 0391 5 208 pages £7.50

Scottish Academic Press
33 Montgomery Street, Edinburgh EH7 5JX.

THE BRITISH JOURNAL OF EDUCATIONAL PSYCHOLOGY

MONOGRAPH SERIES

RECENT ADVANCES IN CLASSROOM RESEARCH

Neville Bennett and Charles Desforges

Research on teaching and learning in classroom contexts has flourished over the last decade and advances have been made both in theory building and in methodology. However along with this increasing research effort there has been a fragmentation in focus orientation and dissemination This has not helped researchers to communicate among themselves or with the practitioner.

The purpose of this Monograph is to aid communication and dissemination of information through reviews of developments in a wide range of research areas by specialists in the field These contributions are grouped into sections each dealing with curriculum content — mathematics and writing, and with context — matching and teacher planning

SBN 7073 0417 2 208 pages £7.50

Scottish Academic Press
33 Montgomery Street, Edinburgh EH7 5JX.

PUBLICATIONS OF THE BRITISH PSYCHOLOGICAL SOCIETY

British Journal of Psychology
British Journal of Medical Psychology
Journal of Occupational Psychology
British Journal of Social Psychology
British Journal of Clinical Psychology
British Journal of Developmental Psychology
British Journal of Mathematical and Statistical Psychology
The British Journal of Educational Psychology

The *British Journal of Psychology*, the *British Journal of Medical Psychology*, *Journal of Occupational Psychology*, the *British Journal of Developmental Psychology*, the *British Journal of Social Psychology* and the *British Journal of Clinical Psychology* appear in quarterly parts The 1985 subscription prices, payable in advance, are £60 00 (US$110 00) for the *British Journal of Psychology*, £38·00 (US$64·50) for the *British Journal of Social Psychology*; £38 00 (US$64·50) for the *British Journal of Clinical Psychology*; £35 00 (US$64·50) for the *British Journal of Developmental Psychology*, £49·00 (US$90·50) for the *British Journal of Medical Psychology*, and £41·00 (US$79·50) for the *Journal of Occupational Psychology*

The *British Journal of Mathematical and Statistical Psychology* is issued twice a year, in May and November. The subscription rate, payable in advance, is £41 00 per volume post free (US$79·50 overseas).

Subscriptions to any of these journals may be sent to a bookseller or subscription agent, or to The British Psychological Society, The Distribution Centre, Blackhorse Road, Letchworth, Herts SG6 1HN, UK. A special price of £273·00 (US$500 00) for all seven journals going to the same address is offered for 1985.

The British Journal of Educational Psychology is issued three times a year, in February, June and November. The prepaid subscription is £18 00 a year, post free, or £6 50 each number. The prepaid subscription price in the USA is $40·00 post free. Subscriptions should be sent to the Scottish Academic Press, 33 Montgomery Street, Edinburgh EH7 5JX.

MANUSCRIPTS

For the *British Journal of Psychology* should be sent to Professor D E. Blackman, Department of Psychology, University College Cardiff, PO Box 78, Cardiff CF1 1XL.

For the *British Journal of Medical Psychology* should be sent to Dr Sidney Crown, Department of Psychiatry, The London Hospital, Whitechapel, E1 1BB.

For the *Journal of Occupational Psychology* should be sent to Dr D Guest, Department of Industrial Relations, London School of Economics and Political Science, Houghton Street, London WC2A 2AE.

For the *British Journal of Social Psychology* should be sent to Dr A. S R Manstead, Department of Psychology, The University, Manchester M13 9PL.

For the *British Journal of Clinical Psychology* should be sent to Dr D Shapiro, MRC/SSRC Social and Applied Psychology Unit, Department of Psychology, The University, Sheffield S10 2TN.

For the *British Journal of Mathematical and Statistical Psychology* should be sent to Dr C. Leach, Department of Psychology, University of Newcastle upon Tyne, Newcastle upon Tyne NE1 7RU.

For the *British Journal of Developmental Psychology* should be sent to Professor P. E. Bryant, Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford OX1 3UD.

For *The British Journal of Educational Psychology* should be sent in duplicate to Professor Hazel Francis, University of London Institute of Education, 20 Bedford Way, London WC1H 0AA

Contributors are advised to consult *Suggestions to Contributors* issued by the Standing Committee on Publications of the British Psychological Society and obtainable, price £2·00 post free, from The British Psychological Society, St. Andrew's House, 48 Princess Road East, Leicester LE1 7DR.

All manuscripts should be submitted in duplicate.

OFFPRINTS

Contributors on application receive fifty copies of their papers free Additional copies may be obtained in units of fifty copies.

REPRINTS

The *Journal* is available in reprint from Swets and Zeitlinger B.V., P.O Box 810, 2160 SZ Lisse, The Netherlands, Vols. 21-29, 31-47.

SUBSCRIPTIONS TO THE BRITISH PSYCHOLOGICAL SOCIETY

The subscription for Fellows, Associates and Graduate Members is £30·00.

They are entitled to purchase any or all of the Society's Journals at the special rates of £6 00 per Journal per annum.

Information on membership and activities may be obtained from the Secretary, The British Psychological Society, St. Andrew's House, 48 Princess Road East, Leicester LE1 7DR.

BRITISH JOURNAL OF EDUCATIONAL PSYCHOLOGY,

VOL. 55, PART 3, 1985

CONTENTS

26 MAY 1986

ARTICLES:

Biggs, J. B. — The Role of Metalearning in Study Processes 185

Boreham, N. C. — Transfer of Training in the Generation of Diagnostic Hypotheses: the Effect of Lowering Fidelity of Simulation 213

Abrams, D., Sparkes, K., and Hogg, M. A. — Gender Salience and Social Identity: the Impact of Sex of Siblings on Educational and Occupational Aspirations 224

Richmond, P. G. — The Relationship of Grade, Sex, Ability and Socio-Economic Status to Parent, Peer and School Affiliation 233

Gill, C. E., Jardine, R., and Martin, N. G. — Further Evidence for Genetic Influences on Educational Achievement 240

Thomas, D. — The Determinants of Teachers' Attitudes to Integrating the Intellectually Handicapped 251

Bigsby, P. — The Nature of Reversible Letter Confusions in Dyslexic and Normal Readers: Misperception or Mislabelling? . 264

Kraft, G. S., Novie, G. J., and Kulhavy, R. W. — Verbal/Spatial Task Expectations in Written Instruction . 273

McKay, M. F., Neale, M. D., and Thompson, G. B. — The Predictive Validity of Bannatyne's WISC Categories for Later Reading Achievement 280

Rusted, J., and Hodgson, S. — Evaluating the Picture Facilitation Effect in Children's Recall of Written Texts 288

Darch, C., and Gersten, R. — The Effects of Teacher Presentation Rate and Praise on LD Students' Oral Reading Performance 295

RESEARCH NOTES:

Lindsay, G., Evans, A., and Jones, B. — Paired Reading Versus Relaxed Reading: A Comparison . 304

Blane, D. C., Pilling, D., and Fogelman, K. — The Use of Longitudinal Data in a Study of Children's School Mobility and Attainment 310

PUBLICATIONS RECEIVED 314

AUTHOR INDEX TO VOLUME 55 317

NOTES TO CONTRIBUTORS 31

CALCUTTA UNIVERSITY CENTRAL LIBRARY

UK ISSN 0007-0998